Mental Health Care Utilization in Prepaid and Fee-for-Service Plans among Depressed Patients in the Medical Outcomes Study

Roland Sturm, Catherine A. Jackson, Lisa S. Meredith, Winnie Yip, Willard G. Manning, William H. Rogers, and Kenneth B. Wells

Objective. We compare mental health utilization in prepaid and fee-for-service plans and analyze selection biases.

Data Source. Primary data were collected every six months over a two-year interval for a panel of depressed patients participating in the Medical Outcomes Study, an observational study of adults in competing systems of care in three urban areas (Boston, Chicago, and Los Angeles).

Study Design. Patients visiting a participating clinician at baseline were screened for depression, followed by a telephone interview, which included the depression section of the NIMH Diagnostic Interview Schedule. Patients with current or past lifetime depressive disorder and those with depressed mood and three other lifetime symptoms were eligible for this analysis. We analyze mental health utilization based on periodic patient self-report.

Analytic Methods. We use two-part models because of the presence of both nonuse and skewness of use. Standard errors are corrected nonparametrically for correlations across observations due to clustered sampling within participating physicians and repeated observations on the same individual.

Principal Findings. The average number of mental health visits was 35–40 percent lower in the prepaid system, adjusted and unadjusted for observed differences in patient characteristics, including health status. Utilization differences were concentrated among patients of psychiatrists, with only minor differences among patients of general medical providers. Analyzing the effect of switches that patients make between payment systems over time, we found some evidence of adverse selection into fee-for-service plans based on baseline utilization, but not based on utilization at the end of the study. In particular, after adjusting for observed patient characteristics and health status, patients switching out of prepaid plans had higher baseline use than predicted, whereas patients switching out of fee-for-service had lower use than predicted. Switching itself appears to be related to an immediate decline in utilization and was not followed by an increase or "catch-up" effect.

Conclusions. The absence of the commonly found "catch-up" effect following switching and the significant decrease in utilization during the switching period

suggests an interruption in care that does not occur for patients staying within a payment system. This finding emphasizes the need for integrating new patients quickly into a system, an issue that should not be neglected in the current policy discussion.

Key Words. Depression, mental health care, utilization, payment system

Many health care reform plans propose relying more on prepaid payment structures. Thus it is important to understand how service use differs between prepaid and fee-for-service plans, especially for the sickest patients, who have the greatest need for care. This article focuses on depression, a particularly important condition because of its prevalence, and its impact on both the individual and society, and because mental health care is perceived to be more responsive to economic incentives than medical care. For the individual, depression is associated with deterioration in functioning that is comparable or worse than the effects of many chronic conditions, including arthritis, diabetes, and hypertension (Wells, Stewart, Hays, et al. 1989). Depression is very common in the community and in practice settings, and is associated with high rates of service utilization and therefore high costs (Regier, Narrow, Rae, et al. 1993; Broadhead et al. 1990). For society, depression is costly and the annual social costs of affective disorders, estimated to be around $44 billion annually, exceed the social costs of coronary heart disease or arthritis (Greenberg et al. 1993). Depression also imposes costs for a longer period of time than other major diseases and imposes a particular burden on employers because it affects relatively young individuals (Greenberg, Stiglin, Finkelstein, et al. 1993).

In this article, we use data from the Medical Outcomes Study (MOS) to examine differences in mental health care utilization for patients receiving prepaid or fee-for-service care. The longitudinal design permits us to study selection over time as patients switched between prepaid and fee-for-service plans, and how these switches were related to service utilization. Patient sickness and need for care were assessed independently from the provider, and this is a major advantage over claims data or record abstractions.

Prior research has shown that some prepaid forms of health care financing incur lower utilization/expenditures than fee-for-service plans for outpatient mental health care (Diehr, Williams, and Martin 1984; Wells, Manning, and Benjamin 1986; Norquist and Wells 1991). Most of these savings were achieved through less intensive styles of care (fewer visits for patients with at least one visit), but similar access to care (similar probability of having any visit). However, these results for populations enrolled in a particular plan or for community populations may not hold for the subpopulation of sick patients. For example, if the prepaid sector eliminates only unnecessary or discretionary visits for healthy individuals, the difference between fee-for-service and prepaid may be smaller or nonexistent for sicker patients. This issue cannot be appropriately addressed without an independent assessment of patient sickness. Moreover, prior studies were largely limited to one or a small number of HMOs, or to one geographic area, or both, thus limiting their representativeness, whereas the MOS sampled a much wider range of prepaid and fee-for-service plans. Prepaid plans in the MOS include both staff model HMOs and independent practice associations (IPAs), which are a more recent, fast growing form of prepaid care. The central question for this article is:

Is there a difference in the utilization patterns among depressed patients, both descriptively and after adjusting for health status and other sociodemographics, between prepaid and fee-for-service plans?

One potential difficulty in observational studies is that observed utilization (and expenditure) differences between prepaid and fee-for-service plans can be affected by biased selection (Luft 1981; Berki and Ashcraft 1980). Such selection effects can be favorable to a payment system if persons using few services enroll in it, or the selection may be adverse if persons using many services select into the system. With substantial biased selection, the observed payment type effects may differ from the "pure" effect due to organizational incentives. It is important to examine closely the possible selection effects given the relevance of nonexperimental data for comparison of system performance and expenditure. In contrast to common beliefs,

switching payment systems is quite common even among depressed patients with an established provider relationship, over 11 percent annually, and it seems to be prone to selection effects (Sturm, McGlynn, Meredith, et al. 1994). In addition, switches between payment systems may have direct effects on a patient's health care. While switches may entail disruptions in established relationships, they may also allow a patient to find more satisfying health care. Health care reform may give more people the opportunity to choose between plans, and this may raise the overall incidence of switching. There have been numerous studies of patient selection in general and of mental health (Luft and Miller 1988; Hellinger 1987; Wilensky and Rossiter 1986), but our longitudinal data permit a more comprehensive analysis. Specifically, we consider the following four questions:

   1. Based on baseline utilization, is there evidence of adverse selection among patients switching payment systems over time? This is the analysis used most often to detect selection bias, typically under the assumption that past utilization will accurately predict future utilization and often because there are no measures of health status. While many studies were limited to comparing fee-for-service stayers with individuals switching into prepaid HMOs, we examine utilization before a patient switches (preswitching utilization) for both switchers into the prepaid and the fee-for-service sector, controlling for initial physical and psychological sickness.

   2. Based on utilization after patients have switched, is there evidence of adverse selection? A major concern among insurers and prepaid health plans is that persons may defer utilization until they enroll in a new plan with broader services and lower out-of-pocket expenses. Maternity care, in particular, has been found to be susceptible to such "stored-up" utilization (Hudes et al. 1980; Robinson, Gardner, and Luft et al. 1993). Such increases are also likely if switching is related to unsatisfied needs. We therefore want to compare whether patients switching payment systems during the study differ in their utilization pattern from similar patients in the same system who have not switched. This is also an important complement to the preceding question, because some researchers suggested that an analysis based on prior utilization overestimates the effect of selection bias (Buchanan and Cretin 1986).

   3. Is there an immediate effect of switching on utilization? Switching payment systems may entail disruptions in the continuity of care, which may be of clinical importance for sick patients. We specifically focus on the six-month period in which a patient switches payment systems.

4. Are utilization differences, switching, and provider specialty related? An important element of prepaid care is its reliance on primary care providers and their role as gatekeepers. Because specialty providers differ in their propensity to treat depression (and in their style of treatment), an analysis of utilization and selection would be incomplete without considering differences among provider specialties (Wells, Hays, Burnham, et al. 1989). If one specialty in a payment system provides substantially fewer services than the same specialty under an alternative payment system, we would expect that patients treated by that specialty are more likely to switch into the alternative system, especially if access to the mental health specialty is restricted in some systems.

METHODS

STUDY DESIGN AND DATA

The Medical Outcomes Study (MOS) is an observational study of adults in competing systems of care in three urban areas (Boston, Chicago, and Los Angeles). In each geographic area, clinicians in HMOs, one or more large multispecialty group practices, and solo/small group practices were enrolled in the study. During the initial cross-sectional data collection, patients of the enrolled providers were screened for depression, myocardial infarction, congestive heart failure, hypertension, and diabetes. Subsequent interviews and examinations were performed to confirm diagnoses and to enroll patients with one or more of these five chronic conditions into the longitudinal part of the study. The screener for depression was a brief self-administered instrument (Burnam et al. 1988). For patients exceeding a cutoff score, the screener was followed by a telephone interview that included the depression section of the NIMH Diagnostic Interview Schedule (Robins et al. 1981). The telephone interview was limited to respondents who had an ongoing relationship with an MOS clinician, could complete self-administered questionnaires, and did not have specific acute physical conditions, such as recent surgery, that severely limited functioning. Fuller discussions of study design, sampling process, and response rates have been published elsewhere (Rogers, McGlynn, Berry, et al. 1992; Rogers, Wells, and Meredith 1993; Tarlov, Ware, Greenfield, et al. 1989).

For this study, we have selected patients in the longitudinal portion of the MOS who had current depressive disorder (major depression or dysthymia), past lifetime (but not current) depressive disorder, or depressive

symptoms (but just missing the criteria for lifetime depressive disorder, i.e., depressed mood and three other lifetime symptoms). Individuals with current depressive disorder met three conditions: (1) DSM-III criteria for a lifetime diagnosis of major depression or dysthymia; (2) an episode of depression or dysthymia during the last 12 months; and (3) no remission from major depression since the onset of the recent episode. We excluded patients with lifetime mania and patients without insurance, leaving a total of 773 individuals with complete data at baseline. All patients in the longitudinal study had an established relationship with an MOS clinician at study baseline.

Self-administered mailed questionnaires (at months 0, 6, 12, 18, and 24 of the study) and telephone-administered interviews (at baseline and at the end of years 1 and 2) are the main sources of information. For depressed patients, there was no variation by type of payment and no interaction effect between payment system and level of depression in the response rates. However, the overall return rates for depressed patients on the mailed questionnaires were substantially lower than those of the medical patients in the longitudinal part of the MOS (Rogers, McGlynn, Berry, et al. 1992).

Independent explanatory variables included sociodemographic information (age, education, ethnicity, gender, marital status) and study site, obtained from baseline questionnaires. In addition, summary measures of psychological and physical illness were developed for each patient using baseline data (Rogers, Wells, Meredith, et al. 1993).

The central dependent variable for this analysis is the number of mental health visits in a six-month interval. A mental health visit is defined as any visit reported to be for personal or emotional problems to either a mental health specialist or a general medical provider, excluding overnight hospital stays. For analytic purposes, we partition mental health visits into two parts: (1) a dichotomous (0-1) variable of whether or not there have been any mental health visits (probability of care), and (2) the number of mental health visits (intensity of care) for patients with one or more visits.

STATISTICAL METHODS

Because of the presence of both nonuse and skewness of use, conventional multiple regression methods are problematic (Duan et al. 1983). We therefore rely on a "two-part" model of the demand for mental health care. The first part is a logit equation for the probability of any outpatient mental health care. The second part regresses the natural logarithm of the number of mental health visits on explanatory variables to analyze the level of use

for patients with one or more visits. The first equation separates users from nonusers and addresses the large number of zero use. The logarithmic transformation of the number of visits for users in the second regression equation alleviates the skewness displayed by the data. A detailed discussion of multipart models can be found in Duan et al. (1983); these models have been used to analyze outpatient mental health by several authors (Wells, Manning, and Benjamin 1986; Horgan 1986; Taube, Kessler, and Burns 1986).

   To predict the number of visits for patients with any use, the predicted dependent variable (Log[number of mental health visits]) was retransformed. We used the nonparametric smearing estimator of Duan (1983), which remains consistent under non-normality when a simple parametric retransformation based on the assumption of normality gives biased predictions. The MOS is a clustered sample design in which many patients from the same provider and clinic are sampled. Moreover, there are repeated observations on the same patient. If these observations are not independent (for example, due to unobserved patient characteristics), standard errors are biased–typically underestimated–and this could lead to spurious significant findings. These correlations across observations do not affect the consistency of parameter estimates, but the variance estimates need to be corrected. This can be done easily, using a technique that is known as "robust standard errors," "sandwich variance estimator," or "Huber correction" (Huber 1967; Schluchter 1988; Liang and Zeger 1986; Neuhaus 1992; White 1982). We performed a number of preliminary specifications tests for the functional form of the model and for the inclusion of covariates using split-sample cross-validation to optimize the predictions from the model (Duan et al. 1983).

   The MOS used a multistage sampling design. We therefore weighted observations for clinician and patient sampling probabilities and the inverse of the recency of the last visit prior to the screening visit for predictions and adjusted numbers. This weighting makes the estimates representative for the population of depressed patients treated in the prepaid and FFS plans included in the study.

   Because prepaid and fee-for-service plans differ in the degree to which they rely on mental health specialty care (Sturm, Meredith, and Wells 1994), we did not include provider specialty in the two-part model because this would lead to incorrect inferences about the overall effect of payment type; the model should control for patient characteristics, but not for organizational characteristics. For example, if the only difference between two

health care systems were in the use of specialists, correcting for specialty would lead us to conclude that there is no difference between systems. Because of the importance of provider specialty for individual care, we analyze differences between the prepaid and the fee-for-service sectors by specialty descriptively in a separate subsection, in which we distinguish general medical providers, psychiatrists, and nonphysician mental health specialists (psychologists, psychiatric nurses).

To study the effect on utilization of switching between payment types, we consider whether switching has any additional explanatory power by analyzing the residual (actual number of mental health visits minus predicted number based on the two-part model). This approach detects systematic selection effects beyond those related to observable differences in patient characteristics (including psychological and physical illness).

RESULTS

Table 1 provides weighted means and standard deviations for the independent variables measured at baseline. The last column in Table 1 reports the t-statistic for testing the hypothesis that the means in fee-for-service (FFS) and prepaid (PP) plans are the same; a negative value indicates a higher mean in the prepaid sector. There are significantly fewer men and nonwhite persons and an indication of lower levels of education among those initially in fee-for-service care, but there are no major differences in age, marital status, psychological health status, or physical health status between fee-for-service and prepaid systems.

REGRESSION ANALYSIS

Parameter estimates for the two-part model to study the effect of observed variables on mental health outpatient utilization are provided in Table 2. Interactions of observed variables with payment type and variables indicating study site were consistently insignificant in preliminary specifications tests and were therefore deleted. Thus, there is no evidence that outpatient mental health utilization by depressed persons differs by geographic location or that the difference between prepaid and fee-for-service plans varies by other patient characteristics.

The probability of any visit, controlling for sociodemographic and health status variables, was significantly lower in prepaid plans ($p < .01$), corresponding to a 6 percentage point reduction in prepaid care (adjusting

Table 1:   Weighted Means and Standard Deviations of
Covariates at Baseline

Variable	FFS (s.d.)	PP (s.d.)	t-*Test for Equality* FFS–PP
Age of patient	43.10 (14.31)	42.20 (13.58)	0.583
Male	0.25 (0.43)	0.40 (0.49)	−3.08***
Nonwhite	0.16 (0.37)	0.29 (0.45)	−2.77***
Married	0.50 (0.50)	0.47 (0.50)	0.684
Education in years	12.86 (3.24)	13.56 (2.72)	−1.76*
MH (Continuous composite measure of psychological health; higher value is sicker)	0.32 (0.77)	0.34 (0.72)	−0.32
PH (Continuous composite measure of physical health; higher value is sicker)	−0.11 (1.06)	−0.06 (1.13)	−0.47
Number of patients	384	389	

$*p \leq .10; **p \leq .05; ***p \leq .01.$

for other covariates, we predict a .56 probability of any mental health visit in fee-for-service and a .50 probability in prepaid plans per six months). We also found a significant difference in the intensity of use for patients with any mental health care utilization ($p < .01$): The estimated difference for users, adjusted for sickness and sociodemographic differences between payment plans, is slightly under five visits per six month period (14 in fee-for-service versus 9 in prepaid). The difference in intensity, rather than the probability of use, is the main reason for the substantial overall difference in utilization between the two payment systems. Combining probability of use and intensity of use and adjusting for sickness and demographic differences, the predicted number of mental health visits per six months is about nine in fee-for-service and 5.5 in prepaid. Interestingly, this is almost identical to the simple unadjusted means. Thus, differences in observed characteristics of the enrolled population are not responsible for utilization differences between prepaid and fee-for-service plans even in unadjusted comparisons.

The initial level of psychological sickness (MH) and its squared term are jointly significant at $p < .01$; a 1 percent increase in psychological

Table 2:   Logit and Linear Regressions for Probability and Level of Mental Health Care

Dependent Variable (Type of Regression)	Any Mental Health Visits (Logit)	ln (Number of Mental Health Visits) If There Were 1 or More Visits (Linear)
Age/10	0.647** (0.320)	0.553*** (0.212)
Age • Age/100	−0.070** (0.034)	−0.066*** (0.022)
Married	−0.190 (0.137)	−0.192** (0.088)
Male	−0.282* (0.156)	−0.200** (0.100)
Nonwhite	−0.660*** (0.175)	−0.200* (0.110)
Education	0.088*** (0.025)	0.063*** (0.016)
MH	1.029*** (0.128)	0.151 (0.113)
MH • MH	−0.010 (0.098)	0.184*** (0.064)
PH	−0.129 (0.062)	−0.093** (0.039)
PPD	−0.335*** (0.127)	−0.383*** (0.081)
Year	−0.489*** (0.058)	−0.123*** (0.039)
Constant	−1.73 (0.77)	0.349 (0.536)
Number of Observations	2712	1528
(Pseudo-)R^2	.1423	.1973

Note: Standard errors in parentheses. Standard errors are corrected for clustered sampling using Huber's method.

*$p \leq .10$; **$p \leq .05$; ***$p \leq .01$.

sickness increased total utilization by 6 percent and most of this effect is due to the number of visits for users rather than through increases in the probability of any visit. Worse physical health (higher PH), controlling for mental health, reduced the probability and level of mental health visits, although the effect is small (a 1 percent increase in physical sickness decreases mental

health visits by about 0.7 percent). However, worse physical health increased the total number of medical and mental health visits (results not reported).

Regarding the other regressor variables, our results are consistent with previous research on mental health care utilization. Age and Age-squared indicate the usual inverted ∪-shaped relationship between age and the use of mental health care for depression, estimated to peak between 40 and 50 years. Married patients, men, and ethnic minorities have lower probabilities and levels of use. Better-educated patients have higher probabilities and levels of use, and the elasticities are very high: the model implies that a 1 percent increase in education increases total utilization by 1.3 percent, of which 2/3 are due to an increase in the number of visits for users. A result specific to the MOS is the decrease in utilization over the course of the study. This is a consequence of the study design, which sampled patients receiving treatment at the beginning of the study.

SWITCHES BETWEEN PAYMENT SYSTEMS AND UTILIZATION

We consider interactions between service use and switching at three points in time: at baseline (before switching), at the end of the study (after switching), and during the switching period.

Table 3 summarizes the differences in utilization at baseline between those who subsequently switch and those who stay within each of the two payment systems; Table 4 gives the number of patients by specialty and switching status. Patients who later switched from fee-for-service to prepaid had significantly fewer mental health visits than patients who stayed in fee-for-service ($p < .01$); the opposite pattern is seen for patients initially in prepaid and switching to fee-for-service (not significant). Of these differences between stayers and switchers in terms of actual visits, only about half can be explained by observable characteristics, including psychological and physical health (compare actual to predicted utilization). In other words, fee-for-service patients eventually switching to prepaid used fewer services at baseline than one would expect given their health status and sociodemographic characteristics, relative to those of all patients.

If the differences between actual and predicted use are the consequence of permanent individual-specific characteristics related to health care, this can have substantial long-run effects on observed payment system differences. In that case, a depressed patient who switched from a fee-for-service plan to a prepaid plan is estimated to have only about 2/3 of the service use that a similar patient would have who has been in the

Table 3:   Unadjusted and Adjusted (Predicted) Baseline Utilization:
Are Switchers Different?

Variable	System at Baseline	Stay in System		Switch Systems	
		Actual	Predicted	Actual	Predicted
Number of mental health visits	FFS***	11	12	6	9
	PP	7	7	10	9

Note: Predicted numbers adjust for patient characteristics and health, based on the two-part model of Table 2.

***Difference between stayers and switchers significant at $p < .01$.

Table 4:   Baseline Utilization by Provider Specialty (Unadjusted
Number of Mental Health Visits)

Payment Type Initial-Year 2	Psychiatrist	General Medical Provider	Other Mental Health Specialist
FFS–FFS	19.9 (58)	4.3 (127)	17.8 (40)
FFS–PP	20.0 (7)	1.3 (28)	5.0 (6)
PP–PP	12.4 (26)	4.5 (141)	12.6 (47)
PP–FFS	7.3 (6)	6.6 (26)	11.3 (11)

Note: Mental health visits during preceding six-month period; number of observations in parentheses; unadjusted, weighted.

prepaid sector for a longer time (six actual visits versus nine predicted visits; see Table 3), whereas a new fee-for-service patient uses 11 percent more services than a fee-for-service stayer. Statistically, these effects are significant: whether or not an individual changed payment types during the course of the study explained a significant fraction of the baseline utilization not captured by observed variables ($p < .05$).

By the end of the study, the differences between switchers and stayers have largely disappeared, and variables indicating whether or not a patient switched during the course of the study are only marginally significant ($p < .10$) (result not shown in a table). Thus, while analyzing utilization before the switch suggests selection bias, the bias is not strong when analyzing utilization after the switch. The most important finding from analyzing postswitch utilization is that switchers from either system had lower utilization after switching than comparable patients not switching, suggesting no evidence of "stored-up" demand and higher "start-up" utilization.

For switchers, we also compared changes in utilization over time to consider the immediate effect of switching. We found that switching was associated with a significant reduction in utilization ($p < .05$), three visits during the six-month period in which patients changed payment systems, but there was no difference by initial payment system. This may indicate that switching is associated with interruptions in care.

THE ROLE OF PROVIDER SPECIALTY

Table 4 reports unadjusted utilization by provider specialty at baseline and whether or not a patient switched payment types during the two years of the study. While there was little difference in use between prepaid and fee-for-service patients of general medical providers, prepaid patients of psychiatrists had significantly fewer visits than fee-for-service patients of psychiatrists ($p < .01$), even though there is no difference in psychological sickness. Even by the end of year 2, patients initially treated by psychiatrists in the prepaid sector had half as many visits as patients initially treated by psychiatrists in fee-for-service (8.5 versus 19.4 visits). The difference in use between prepaid and fee-for-service patients of other mental health specialists lay between psychiatrists and general medical providers but was not statistically significant.

Unfortunately, due to payment type and provider switching, we cannot reliably determine the specialty of the main provider at the end of the study for the sample. However, the difference in utilization between prepaid and fee-for-service patients of psychiatrists is as strong for those patients who neither switched payment types nor individual providers throughout the study (13.5 in prepaid versus 20.3 visits in fee-for-service). This latter comparison is based on only 59 patients but remains marginally significant ($p < .10$).

The second main finding regarding provider specialty is that, within each specialty, switchers had lower baseline utilization than patients staying in the system. This finding is mainly due to the fee-for-service sector, where the difference between switchers and stayers was significant for both general medical providers ($p < .05$) and other mental health specialists ($p < .01$). The preceding finding–that patients switching out of prepaid were higher users overall than prepaid stayers–is related to differences in the specialty mix: patients of mental health specialists were more likely to leave prepaid plans than patients of general medical providers, whereas patients of general medical providers were more likely to leave fee-for-service plans (Sturm, McGlynn, Meredith, et al. 1994). Thus, the finding of Table 3 is explained

by the relatively high proportion of patients of psychiatrists who leave the prepaid system.

DISCUSSION

Regarding the central question of this study, whether there are differences in utilization patterns between prepaid and fee-for-service plans, we find that depressed prepaid patients obtained substantially fewer mental health services than similar patients in fee-for-service care. Self-reported information on coverage suggests that this difference is not due to more generous coverage of mental health care in fee-for-service plans, implying that the difference is more likely a supply- rather than a demand-side effect. The utilization difference between fee-for-service and prepaid plans for depressed patients in the MOS was comparable to previously reported differences in outpatient mental health care for enrolled populations (Wells, Manning, and Benjamin 1986) and amounts to approximately 700 additional mental health visits per 100 depressed patients in the fee-for-service sector. There were no differences in the level of psychological or physical sickness between prepaid and fee-for-service plans. We also compared mental health hospitalizations, but there are relatively few observations, limiting the statistical power and precision of our analysis. It appears that differences in hospitalization rates for mental health reasons account for between one and three additional stays annually in fee-for-service plans per 100 patients. In contrast to other clinical conditions, including other severe mental illnesses like schizophrenia, cost differences between prepaid and fee-for-service plans seem to be largely due to differences in outpatient utilization.

Consistent with earlier research (Wells, Manning, and Benjamin 1986), our study found that most of the difference between prepaid and fee-for-service came from fewer visits (lower intensity) for patients who had at least one visit, rather than from a difference in the proportion of patients having any use. However, the difference between payment systems regarding the probability of any mental health visit was statistically significant. In both payment systems, male and nonwhite patients were significantly lower users of outpatient mental health services and more educated and psychologically sicker patients were higher users. Although the interaction between ethnicity and payment type in the probability of care was significant by itself (nonwhite patients were relatively more likely to receive any care in prepaid), no such effect existed for the intensity of care, and the overall ethnicity-payment

interaction was insignificant. Nevertheless, a relatively higher access to care (in terms of probability of any mental health visit) for nonwhite patients in prepaid care is in agreement with our earlier finding that nonwhite patients were unlikely to leave prepaid plans (Sturm, McGlynn, Meredith, et al., 1994).

   While the two-part regression framework controls for selection bias according to observed variables, including health, the possibility remains of unobserved heterogeneity among patients by payment type (selection bias that cannot be controlled for by observed variables). This brings us to the specific questions regarding the relationship of utilization and switches between payment systems.

   Our first question was: **Based on baseline utilization, is there evidence of adverse selection among patients switching payment systems during the course of the study?** We found that patients who switched from fee-for-service to prepaid plans were among the lowest users of services while they were enrolled in the fee-for-service sector. The pattern is opposite for patients leaving prepaid plans, who were among the highest users before switching. About half of this selection effect is explained by observed variables at baseline, including physical and psychological health, indicating an important role for patient preferences. Patients switching out of prepaid may have preferred higher levels of care than stayers (indicated by the fact that prepaid stayers had lower and switchers higher baseline utilization than predicted by observed patient characteristics and health status) and therefore desired more services than prepaid systems considered necessary and were willing to supply.

   The effect of unobserved heterogeneity could be quite substantial. Under the assumption that the differences between actual and predicted utilization reflect permanent individual-specific characteristics related to health care (such as the propensity to seek intensive mental health care), the amount of switching over the two years of the study in this sample could imply that the level of utilization in the prepaid sector is only 90 percent of the level in the fee-for-service sector by the end of the study. This 10 percent reduction would be solely due to unobserved heterogeneity and the ensuing adverse selection and is one-third of the total estimated difference between prepaid and fee-for-service plans. The selection effect could be even higher because of selection before the beginning of the study.

   Since some researchers suggested that analyzing utilization prior to a switch can overstate selection biases, we considered the second question: **Based on utilization in year 2 of the study, is there evidence of adverse**

selection? While the preceding discussion suggests a very substantial selection effect, most of the differences between switchers and stayers have disappeared by the end of the study. This provides some evidence that unobserved heterogeneity does not constitute a permanent effect on utilization and that the selection effect discussed above constitutes an upper bound. Because of a general decrease in utilization over the course of the study, which reduced the statistical power (the typical "regression to the mean" for sick patients; Welch 1985), it is difficult to assess how important the baseline selection effect was. We tried various statistical tests to determine whether unobserved heterogeneity differences were permanent or transient, but our results were inconclusive. It is unfortunate that the size of the MOS depression sample is too small to provide a definite answer. Nevertheless, our results strongly suggest that studies analyzing only pre- or postswitching utilization are incomplete.

An important finding in the analysis of this second question was the absence of a stored-up demand effect; that is, patients did not increase their utilization after switching payment plans (even after adjusting for the expected decrease in utilization over time). This differs from the disenrollment studies, reviewed in Luft and Miller (1988), which suggested that HMO disenrollees had higher health care use after disenrollment than comparable fee-for-service enrollees did. The focus on depression may contribute to our different result. It is possible that a patient with mental health problems finds it more difficult to reconnect with a new health care system than a patient with general medical conditions. Because patients were sampled during an acute phase of their depression and their mental health status improved over time, they may have found it unnecessary to expend the effort to obtain access to care. But long-run maintenance therapies can be useful for depression (Kupfer, Frank, Perel, et al., 1992) and this strengthens the argument to make access to mental health care easy for new enrollees. We find this decrease in utilization even though some patients switching payment systems maintain a relationship with their original provider and those patients do not need to establish a new relationship. The prevalence of such "rollovers" is likely to increase with the growth of IPAs.

Analyzing our third question, **Is there an immediate effect of switching on utilization?**, we found a result even more at variance with previous research on selection. Instead of an increase after switching, we find that the most immediate effect of switching payment systems is a significant fall in utilization during the period a patient switches. This decrease could be a consequence of interrupting an established patient-provider relationship

combined with difficulties in accessing mental health care in a new payment system. Although our data do not allow us to study the issue in more detail, it appears that this is a particular difficulty for depressed patients leaving prepaid plans: these patients had higher service use at baseline than predicted (after adjusting for payment type, patient characteristics, and health status), but lower service use at the end of the study than predicted. Again, our different study population is likely to be responsible for this finding: all patients in the depression panel of the MOS had an established provider relationship at baseline. In contrast, previous work focused on enrolled populations and most switchers had no established relationship with a provider. This highlights once more the importance of focusing specifically on sick patients, who are most directly affected by health care reform and for whom the conventional wisdom based on population studies may not apply.

Our final question asked: **Are utilization differences and switching related to provider specialty?** The main finding is that prepaid patients of psychiatrists had substantially fewer visits than fee-for-service patients of psychiatrists, whereas no such difference exists among patients of general medical providers. Other mental health specialists fall in between, but differences are not statistically significant. This mirrors the previously reported outcomes results from the MOS (Rogers, Wells, Meredith, et al. 1993): prepaid patients of psychiatrists had significantly worse functioning outcomes than fee-for-service patients of psychiatrists, but there were no differences in outcomes between fee-for-service and prepaid plans for patients of general medical providers. Thus, there may be a process-outcome link between a less intensive treatment style and worse functioning outcomes. Reduced utilization may lead to less monitoring of the effectiveness of treatment, possibly a weaker relationship between patient and provider and more provider switching, and adverse outcomes. For example, prepaid patients of psychiatrists were more likely to end this relationship earlier than fee-for-service patients of psychiatrists, which in turn was associated with an effective antidepressant medication therapy (Sturm, Meredith, and Wells 1994). However, an observational study such as the MOS cannot determine whether utilization is the cause or the consequence of these other dimensions in which prepaid and fee-for-service care differ.

Provider specialty also explains why patients switching out of prepaid had higher baseline utilization than prepaid stayers, and why patients leaving fee-for-service had lower utilization than stayers: prepaid patients of mental health specialists (both psychiatrists and psychologists) were

significantly more likely to switch than fee-for-service patients of mental health specialists, whereas fee-for-service patients of general medical providers were more likely to switch than prepaid patients of general medical providers (Sturm, McGlynn, Meredith, et al. 1994). Controlling for health status does not affect this utilization difference between stayers and switchers in the two payment systems. In other words, patients leaving prepaid plans have higher use than one would expect given their health status, which is caused by a higher proportion of mental health specialty care than expected according to their health status—and which ultimately reflects the propensity of these patients' to seek care, as well as other unobserved "tastes."

LIMITATIONS OF THE STUDY

The utilization data in this study come from self-reports from the periodic mailed surveys. This allows us to examine utilization irrespective of payment system, which is not possible in claims-based information. The disadvantage is a potential reporting bias. Patients were asked to indicate the number of visits to mental health specialists and the number of visits to non-mental health specialists for mental health or emotional reasons. Thus, the analysis relies on the accuracy of identifying and recording visits for mental health reasons. However, the long recall interval (six months) may lead patients to misestimate utilization. If high users underreport utilization, the study may have underestimated outpatient utilization and the difference between prepaid and fee-for-service plans.

The study design imposes further limits on the scope and generalizability of our findings. One limitation for the analysis is the relatively short duration (two years) of the MOS and the lack of information about the length of time in the baseline health plan or the type of payment plan used prior to the study. Consequently, our estimates are incomplete because we do not have information on switches preceding the beginning of the study and we treat patients switching prior to baseline as stayers. For the comparison of Table 3, this data limitation is likely to underestimate the difference between movers and stayers.

Second, the absence of details about particular plans does not allow us to distinguish differences between plans within each payment system and switches between plans within a payment system. For example, we cannot reliably distinguish whether a patient switches into a staff model HMO or an IPA. Nevertheless, staff model HMOs and IPAs have similar financial incentives to deliver care, which makes it important to distinguish these types of plans from fee-for-service plans.

Finally, the MOS is geographically and temporally limited, and changes in levels of coverage or managed care procedures in both prepaid and fee-for-service plans since the late 1980s may have altered utilization patterns. However, the MOS sampled 15 group practice–style HMOs and IPAs in three geographic locations and the findings may therefore be more representative of the typical differences between the prepaid and the fee-for-service financing systems than previous studies based on only one HMO or geographic site.

CONCLUSION

Despite its limitations, we believe that this study provides a number of new insights that remain relevant in the current policy debate. First of all, previously reported differences in utilization for general populations between payment systems hold for mental health care utilization by depressed patients, even after adjusting for health status. Moreover, these differences in utilization tend to be concentrated among patients of mental health specialists, with only minor differences among patients of general medical providers. This may partly be because rates of use were low in the general medical sector, leaving little room for reductions. Our investigation of payment switching showed a substantial adverse selection effect for fee-for-service using preswitching utilization, but this effect disappeared over the course of the study. While we cannot provide a final answer to the question of whether this biases the estimated differences between fee-for-service and prepaid plans, it raises a problem that future observational studies have to take into account.

We also found two effects that differ from findings for enrolled populations, indicating the importance of studying populations most directly affected by changes in health care systems. The first finding is an absence of increased utilization following a switch. This may indicate patients' difficulties to reconnect in a new mental health system. Even more important, we found a significant decrease in utilization during the period in which a patient switched systems, suggesting possible interruption in care and emphasizing the need for quickly integrating patients into a new system. Unfortunately, patients with major psychiatric disorders, such as depression, may lag behind patients with physical health problems on this account. For example, primary care providers detect depression in only half of all depressed patients (Wells, Hays, Burnham, et al. 1989), implying that a

large fraction of those new to a system that relies on primary care providers as gatekeepers may only receive appropriate treatment after a substantial delay, if at all. The integration of new depressed patients in a system may therefore be an important issue that should not be neglected in the current policy discussion.

ACKNOWLEDGMENTS

REFERENCES

Berki, S. E., and M. L. Ashcraft. 1980. "HMO Enrollment: Who Joins What and Why: A Review of the Literature." *Milbank Memorial Fund Quarterly* 58 (4): 588–632.

Broadhead, W. E., D. G. Blazer, L. K. George, and C. K. Tse. 1990. "Depression, Disability Days, and Days Lost from Work in a Prospective Epidemiologic Survey." *Journal of the American Medical Association* 264 (19): 2524–28.

Buchanan, J., and S. Cretin. 1986. "Risk Selection of Families Electing HMO Membership." *Medical Care* 24 (1): 39–51.

Burnam, M. A., K. B. Wells, B. Leake, and J. Landsverk. 1988. "Development of a Brief Screening Instrument for Detecting Depressive Disorders." *Medical Care* 26 (8): 775–89.

Diehr, P., S. J. Williams, and D. P. Martin. 1984. "Ambulatory Mental Health Services Utilization in Three Provider Plans." *Medical Care* 2 (1): 1–13.

Duan, N. 1983. "Smearing Estimate: A Nonparametric Retransformation Method." *Journal of the American Statistical Association* 78 (September): 605–10.

Duan, N., W. G. Manning, C. N. Morris, and J. P. Newhouse. 1983. "A Comparison of Alternative Models for the Demand for Medical Care." *Journal of Business and Economic Statistics* 1 (2): 115–26.

Greenberg, P. E., L. E. Stiglin, S. N. Finkelstein, E. R. Berndt. 1993a. "The Economic Burden of Depression in 1990." *Journal of Clinical Psychiatry* 54 (11): 405–18.

———. 1993b. "Depression: A Neglected Major Illness." *Journal of Clinical Psychiatry* 54 (11): 419–24.

Hellinger, F. J. 1987. "Selection Bias in Health Maintenance Organizations: Analysis of Recent Evidence." *Health Care Financing Review* 9 (Winter): 55.

Horgan, C. M. 1986. "The Demand for Ambulatory Mental Health Services from Specialty Providers." *HSR: Health Services Research* 21 (2): 291–320.

Huber, P. J. 1967. "The Behavior of Maximum Likelihood Estimates under Nonstandard Conditions." *Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1: 221–33.

Hudes, J., C. A. Young, L. Sohrab, and C. N. Trinh. 1980. "Are HMO Enrollees Being Attracted by a Liberal Maternity Benefit?" *Medical Care* 18 (6): 635–48.

Kupfer, D. J., E. Frank, J. M. Perel, C. Cornes, A. G. Mallinger, M. E. Thase, A. B. McEachran, and V. J. Grochonciski. 1992. "Five-Year Outcome for Maintenance Therapies in Recurrent Depression." *Archives of General Psychiatry* 49 (October): 769–73.

Liang, K. Y., and S. L. Zeger. 1986. "Longitudinal Data Analysis Using Generalized Linear Models." *Biometrika* 73 (1): 13–22.

Luft, H. 1981. *Health Maintenance Organizations: Dimensions of Performance.* New York: John Wiley and Sons.

Luft, H. S., and R. H. Miller. 1988. "Patient Selection in a Competitive Health Care System." *Health Affairs* 7 (1): 97–119.

Neuhaus, J. M. 1992. "Statistical Methods for Longitudinal and Clustered Designs with Binary Responses." *Statistical Methods in Medical Research* 1 (3): 249–73.

Norquist, G. S., and K. B. Wells. 1991. "How Do HMOs Reduce Outpatient Mental Health Care Costs?" *American Journal of Psychiatry* 148 (1): 96–101.

Regier, D. A., W. E. Narrow, D. S. Rae, R. W. Manderscheid, B. Z. Locke, and F. K. Goodwin. 1993. "The De Facto U.S. Mental and Addictive Disorders Service System." *Archives of General Psychiatry* 50 (1): 85–94.

Robins, L. N., J. E. Helzer, J. Croughan, and K. S. Ratcliff. 1981. "National Institute of Mental Health Diagnostic Interview Schedule: Its History, Characteristics, and Validity." *Archives of General Psychiatry* 38 (3): 381–89.

Robinson, J. C., L. B. Gardner, and H. S. Luft. 1993. "Health Plan Switching in Anticipation of Increased Medical Care Utilization." *Medical Care* 31 (1): 43–51.

Rogers, W. H., E. A. McGlynn, S. H. Berry, E. G. Nelson, E. Perrin, M. Zubkoff, S. Greenfield, K. B. Wells, A. L. Stewart, S. B. Arnold, and J. E. Ware. 1992. "Methods of Sampling." In *Measuring Functioning and Well-Being: The Medical*

Outcomes Study Approach, edited by A. L. Stewart and J. E. Ware. Durham, NC: Duke University Press.

Rogers, W. H., K. B. Wells, L. S. Meredith, R. Sturm, and A. Burnam. 1993. "Course of Depression for Adult Outpatients under Prepaid or Fee-for-Service Financing." *Archives of General Psychiatry* 50 (7): 517–26.

Schluchter, M. D. 1988. "Analysis of Incomplete Multivariate Data Using Linear Models with Structured Covariance Matrices." *Statistics in Medicine* 7 (1/2): 317–24.

Sturm, R., E. A. McGlynn, L. A. Meredith, K. B. Wells, W. G. Manning, and W. H. Rogers. 1994. "Switches Between Prepaid and Fee-for-Service Health Systems among Depressed Outpatients: Results from the Medical Outcomes Study." *Medical Care* 31 (September): 917–29.

Sturm, R., L. A. Meredith, and K. B. Wells. 1994. "Provider Choice and Continuity for the Treatment of Depression." Contract No. (DRU) 692-AHCPR. RAND: Santa Monica, CA.

Tarlov, A., J. E. Ware, S. Greenfield, E. C. Nelson, E. Perrin, and M. Zubkoff. 1989. "The Medical Outcomes Study: An Application of Methods for Monitoring the Results of Medical Care." *Journal of the American Medical Association* 262 (70): 925–30.

Taube, C. A., L. B. Kessler, and B. J. Burns. 1986. "Estimating the Probability and Level of Ambulatory Mental Health Services Use." *Health Services Research* 21 (2): 321–40.

Welch, W. P. 1985. "Regression toward the Mean in Medical Care Costs: Implications for Biased Selection in Health Maintenance Organizations." *Medical Care* 23 (11): 1234–41.

Wells, K. B., W. G. Manning, and B. Benjamin. 1986. "Use of Outpatient Mental Health Service Plans: Results from a Randomized Controlled Trial." *Health Services Research* 21 (4): 453–74.

Wells, K. B., A. Stewart, R. D. Hays, M. A. Burnam, W. Rogers, M. Daniels, S. Berry, S. Greenfield, and J. E. Ware. 1989. "The Functioning and Well-Being of Depressed Patients: Results from the Medical Outcomes Study." *Journal of the American Medical Association* 262 (7): 914–19.

Wells, K. B., R. D. Hays, M. A. Burnam, W. Rogers, S. Greenfield, and J. E. Ware. 1989. "Detection of Depressive Disorder for Patients Receiving Prepaid or Fee-for-Service Care: Results from the Medical Outcomes Study." *Journal of the American Medical Association* 262 (23): 3298–302.

White, H. J. 1982. "Maximum Likelihood Estimation of Misspecified Models." *Econometrica* 50 (4): 483–99.

Wilensky, G. R., and L. F. Rossiter. 1986. "Patient Selection in HMOs." *Health Affairs* 5 (1): 66–80.