

The Effects of Mental Health Parity on Spending and Utilization for Bipolar, Major Depression, and Adjustment Disorders

Alisa B. Busch, M.D., M.S.

Frank Yoon, Ph.D.

Colleen L. Barry, Ph.D., M.P.P.

Vanessa Azzone, Ph.D.

Sharon-Lise T. Normand, Ph.D.

Howard H. Goldman, M.D., Ph.D.

Haiden A. Huskamp, Ph.D.

Objective: The Mental Health Parity and Addiction Equity Act requires insurance parity for mental health/substance use disorder and general medical services. Previous research found that parity did not increase mental health/substance use disorder spending and lowered out-of-pocket spending. Whether parity's effects differ by diagnosis is unknown. The authors examined this question in the context of parity implementation in the Federal Employees Health Benefits (FEHB) Program.

Method: The authors compared mental health/substance use disorder treatment use and spending before and after parity (2000 and 2002, respectively) for two groups: FEHB enrollees diagnosed in 1999 with bipolar disorder, major depression, or adjustment disorder (N=19,094) and privately insured enrollees unaffected by the policy in a comparison national sample (N=10,521). Separate models were fitted for each diagnostic group. A difference-in-difference design was used to control for secular time trends and to better reflect the specific impact of parity on spending and utilization.

Results: Total spending was unchanged among enrollees with bipolar disorder and major depression but decreased for those with adjustment disorder (−\$62, 99.2% CI=−\$133, −\$11). Out-of-pocket spending decreased for all three groups (bipolar disorder: −\$148, 99.2% CI=−\$217, −\$85; major depression: −\$100, 99.2% CI=−\$123, −\$77; adjustment disorder: −\$68, 99.2% CI=−\$84, −\$54). Total annual utilization (e.g., medication management visits, psychotropic prescriptions, and mental health/substance use disorder hospitalization bed days) remained unchanged across all diagnoses. Annual psychotherapy visits decreased significantly only for individuals with adjustment disorders (−12%, 99.2% CI=−19%, −4%).

Conclusions: Parity implemented under managed care improved financial protection and differentially affected spending and psychotherapy utilization across groups. There was some evidence that resources were preferentially preserved for diagnoses that are typically more severe or chronic and reduced for diagnoses expected to be less so.

(*Am J Psychiatry* 2013; 170:180–187)

The Mental Health Parity and Addiction Equity Act (MHPAEA) (1), implemented in 2010, requires parity in coverage for mental health/substance use disorders and general medical services. By eliminating benefit limits on mental health/substance use disorder services, such as higher copayments and caps on the number of inpatient days and outpatient visits covered by health plans (2), a primary aim of parity advocates was to increase financial protection for individuals with the most disabling conditions. The logic was that those with the most severe conditions used more services and encountered those limits more often. As a result, they were likely to spend more, placing them at risk of greater financial losses.

Previous studies have found that parity did not lead to increases in mental health/substance use disorder service use or total spending, but that beneficiary out-of-pocket costs were significantly lower after implementation

of parity relative to a comparison group of unaffected health plan members (3–6). Those studies examined the effects of a parity directive instituted in the Federal Employees Health Benefits (FEHB) Program in January 2001, and, more recently, the effects of a 2007 Oregon state parity law that is similar to the MHPAEA. However, the studies did not examine possible differential effects of parity based on diagnosis. Because of the recent implementation of the MHPAEA and its regulations, there is no empirical evidence yet on the law's effects. Additionally, because the MHPAEA is implemented nationwide and the Patient Protection and Affordable Care Act extends MHPAEA provisions to new plans offered through state health insurance exchanges and to Medicaid benchmark plans, there is no obvious national comparison group of individuals who will not experience parity that would allow a rigorous evaluation of MHPAEA's effects on individuals

This article is featured in this month's AJP **Audio** and is discussed in an **Editorial** by Dr. McCarty (p. 140)

with relatively severe compared with those with less severe disorders.

Given the lack of a national comparison group to assess whether the MHPAEA's effects might differ by diagnostic severity, we explored this question in the context of the FEHB Program parity directive, a similar (although not identical) parity policy. In June 1999, President Clinton directed the Office of Personnel Management (OPM) to require all FEHB plans to offer comprehensive parity benefits for in-network services as of 2001. The FEHB Program, covering approximately 8.5 million enrollees, is the largest private health insurer in the United States. The OPM encouraged plans to use managed care techniques to control any increases in mental health/substance use disorder expenditures that could result from the parity benefit expansion.

Using a difference-in-difference design comparing a national sample of FEHB Program enrollees with a national comparison group of individuals included in the Truven Health Analytics MarketScan data, we examined the impact of the FEHB Program parity policy on spending and intensity of service use for individuals diagnosed with any of three disorders: bipolar disorder, a typically chronic and severe illness; major depression, a disorder that exhibits greater heterogeneity in both severity and chronicity; and adjustment disorder, a condition expected to be typically acute and less severe. In this study we use the term "parity policy" to refer to the combined effects of benefit design and the management of care, which we cannot disaggregate. Given that this is the same insurance context of the current MHPAEA, empirical evidence regarding the effects of the FEHB Program parity policy can shed light on how the new MHPAEA law might differentially affect treatment for mental illnesses of varying diagnostic severity levels.

Method

Health Plans

For this analysis, we pooled data from the seven FEHB plans studied in the original evaluation of the FEHB Program parity policy (5). The plans were selected on the basis of region, population size, and interest in participation. All were preferred provider organizations (PPOs). Of the seven plans, four contracted with managed behavioral health carve-out organizations to manage mental health/substance use disorder service use both before and after parity implementation, two implemented carve-outs at the same time the parity policy took effect, and one managed service use internally (i.e., no carve-out) both before and after parity implementation. The MarketScan database comparison group included enrollees of PPO health plans operated by large, self-insured employers.

We used administrative data from the FEHB and MarketScan plans that included enrollment and inpatient, outpatient, and pharmacy claims, as well as their associated costs (total and out-of-pocket), for 2 years before parity implementation (1999 and 2000) and compared them with the second year after parity implementation (2002). We focused on the second year anticipating that any effects of parity would be more

prominent after health plans had an opportunity to adjust to the new policy.

Selection of Cohort

To ensure that the results reflected the effects of parity and not changes in plan enrollee composition, we required continuous enrollment for all four study years. Using the baseline year (1999), we divided the study population into mutually exclusive diagnostic cohorts that typically differ in illness severity and/or chronicity: bipolar disorder, major depression, and adjustment disorder.

To be included in a given diagnostic cohort, an individual was required to have, in 1999, at least two claims with the target diagnosis on different service dates, a single inpatient claim with the target diagnosis, or a single outpatient claim with the target diagnosis if there was no more than one other claim with a different diagnosis. Before creating the cohorts, we excluded persons who had a schizophrenia diagnosis (ICD-9 code 295) in 1999. Using the 1999 claims, we then established the bipolar disorder cohort (codes 296.0–296.1, 296.4–296.8, 301.11, 301.13) using the above algorithm. For example, enrollees were in the bipolar disorder cohort if they had no schizophrenia claims in 1999 and had at least two bipolar disorder claims on different service dates in 1999; or if one bipolar disorder claim, it was inpatient; or if outpatient, the single bipolar disorder claim was one of only two total claims for mental health/substance use disorder services. Claims data of persons not included in the bipolar disorder cohort were then examined to establish the major depression cohort (ICD-9 codes 296.2 and 296.3) using the same algorithm. Finally, the adjustment disorder cohort (code 309) was established from the remaining enrollees not yet selected into a cohort, again using the same algorithm.

Outcomes

We examined two types of outcomes: mental health/substance use disorder spending (total and out-of-pocket) and utilization. We defined utilization consistent with a previous study by Goldman et al. (5): annual utilization outcomes included the number of psychotherapy visits, medication management visits, inpatient mental health/substance use disorder days, and mental health/substance use disorder prescription fills. We defined mental health/substance use disorder medications in two ways: medications used only in the treatment of mental health conditions or substance use disorders, and an expanded list that included medications that could be used for other conditions as well (e.g., valproate). The latter counted as mental health/substance use disorder care only if the enrollee used any mental health/substance use disorder services in the same calendar year as the medication.

Explanatory Variables

Models included regional dummy variables (Northeast, West, South, and Midwest) and patient characteristics (sex, employee/dependent status, and age [centered]). We also included an interaction term for sex and employee status because a preliminary examination indicated that there may be an interaction between these two characteristics.

Statistical Analysis

We used a difference-in-difference approach to account for secular trends in spending and utilization during the study period. The primary explanatory variables of interest were an indicator of whether an individual was enrolled in an FEHB plan (versus a comparison group plan), an indicator for study year, and their interaction. The interaction term is the difference-in-difference estimator, measuring the impact of parity on spending and utilization, controlling for secular time trends and thereby reflecting

TABLE 1. Characteristics of Continuously Enrolled FEHB Program Plan and Comparison Plan Enrollees With Bipolar Disorder, Major Depression, or Adjustment Disorder (N=29,615)^a

Characteristic	FEHB Program Plans (N=19,094)		Comparison Plans (N=10,521)	
	Mean	SD	Mean	SD
Age (years) at baseline	46.2	8.3	43.6	11.0
	N	%	N	%
Female	12,392	64.9	7,211	68.5
Employee	11,827	61.9	6,692	63.6
Region				
Northeast	2,177	11.4	1,755	16.7
South	12,271	64.3	1,671	15.9
Midwest	1,292	6.8	6,262	59.5
West	3,354	17.6	833	7.9
Psychiatric diagnosis				
Bipolar disorder	2,557	13.4	1,177	11.2
Major depression	10,412	54.5	5,245	49.9
Adjustment disorder	6,125	32.1	4,099	39.0

^a FEHB Program=Federal Employees Health Benefits Program. We compared enrollees of seven FEHB Program plans with enrollees of comparison plans operated by large, self-insured employers in the Truven Health Analytics MarketScan database over the period 1999–2002.

the impact of parity specifically. Separate models were fitted to each outcome by diagnostic group.

We used a generalized estimating equation approach to model the spending and utilization outcomes (7). The use of generalized estimating equations permits flexible modeling of the patient-level correlations over time. For both spending outcomes, we estimated two-part models (8) to handle zeros in the analysis of the log spending amounts. The first part estimates the parity policy's effect on the probability that any spending occurred in the given year, and the second estimates its effect on the log spending amount for those who had spending; the second part of the model estimates an effect conditional on use. To generalize the policy's effect on the full population, we report the total changes in spending for the whole population. We observed that spending outcomes were highly skewed, given the illness severity of many of our cohort enrollees, and addressed this using a log-transformation. The probability of any mental health/substance use disorder service use was modeled by probit regression, and the total and out-of-pocket log spending amounts by ordinary least squares. Mental health/substance use disorder service utilization was characterized by Poisson counts that measure the number of annual service units per enrollee in the preparity year and the second postparity year.

Because the interaction terms from our models do not have an intuitive interpretation on their original scale, we transformed results to either the dollar (spending) or percent (utilization) scales. Raw dollar amounts were calculated using Duan's smearing estimate (9). All bootstrap methods were implemented to approximate point estimates and confidence intervals. We report results and statistical tests based on 2,000 bootstrap samples.

Modeling the two parts of spending (probability of spending and spending conditional on use) separately may bias the estimates if the correlation between use and spending outcomes is ignored (10). As a check on this, we simultaneously fitted the two-part models with a common subject-level random effect through a Bayesian approach in the R statistical software package (11; www.r-project.org). The results from this approach were qualitatively the same, and we report the findings from the separate models. All other statistical models, including Poisson regression, were estimated by PROC GENMOD in SAS, version 9.2 (SAS Institute, Cary, N.C.).

A risk of testing multiple outcomes involves falsely rejecting a null hypothesis (i.e., no effect of parity), and adjustments for multiplicity are often used. We used a conservative adjustment of the confidence intervals to 99.2% to maintain an overall type I error rate of 5% for the six tests per diagnostic group (12). As

a sensitivity analysis, we also computed the less conservative Hochberg adjustment. The estimates and standard errors for all comparisons, including postparity year 2001 (which is not a focus of our analysis), are provided in the data supplement that accompanies the online edition of this article. Finally, we computed the preparity out-of-pocket spending across all users within each population (adjusted for age, sex, employee status, and geographical region) to facilitate comparison of the spending changes resulting from parity.

Results

FEHB Program and comparison group enrollees were similar with respect to sex, age, employee status, and the proportion of individuals with a given diagnosis (Table 1). The greatest difference between the two groups was that FEHB Program enrollees resided predominantly in the South (FEHB enrollees, 64.3%; comparison enrollees, 15.9%) and comparison cohort enrollees in the Midwest (FEHB enrollees, 6.8%, comparison enrollees, 59.5%).

In both the FEHB Program and comparison groups, the probability of having any mental health/substance use disorder use and the average total and out-of-pocket mental health/substance use disorder spending conditional on use in the subsequent year (2000) were highest for the bipolar disorder group and lowest for the adjustment disorder group (Table 2). Similarly, the probability of using each type of service in 2000 was highest for the bipolar disorder group and lowest for the adjustment disorder group.

All difference-in-difference spending and service utilization estimates before parity and in the second year post-parity among FEHB Program and comparison group enrollees reflect changes across all enrollees in each diagnostic cohort, not changes conditional on service use. Total spending in the FEHB Program was statistically unchanged for enrollees in the bipolar and major depression diagnostic groups but decreased for the adjustment disorder group (Table 3). Out-of-pocket spending declined after implementation of parity for all three diagnostic categories

TABLE 2. Annual Service Utilization and Spending Before Parity Implementation (2000) Among Enrollees Diagnosed in 1999 With Bipolar Disorder, Major Depression, or Adjustment Disorder in FEHB Program Plans and Comparison Plans (N=29,615)^a

Variable	FEHB Program Plans (N=19,094)		Comparison Plans (N=10,521)	
Diagnostic group totals	N		N	
Bipolar disorder	2,557		1,177	
Major depression	10,412		5,245	
Adjustment disorder	6,125		4,099	
Probability of any mental health/substance use disorder service use	N	%	N	%
Bipolar disorder	2,287	89.4	1,104	93.8
Major depression	8,791	84.4	4,699	89.6
Adjustment disorder	3,483	56.9	2,599	63.4
Conditional on use, per-person mental health/substance use disorder spending	Mean	SD	Mean	SD
Bipolar disorder				
Total spending	\$3,116	\$5,293	\$3,576	\$5,220
Out-of-pocket spending	\$787	\$1,086	\$389	\$430
Major depression				
Total spending	\$1,929	\$3,049	\$2,414	\$3,931
Out-of-pocket spending	\$563	\$732	\$301	\$427
Adjustment disorder				
Total spending	\$1,105	\$1,431	\$1,214	\$1,785
Out-of-pocket spending	\$428	\$432	\$179	\$241
Mental health/substance use disorder service use				
Bipolar disorder	N	%	N	%
Any psychotherapy visit	1,328	51.9	725	61.6
Any medication management visit	1,023	40.0	570	48.4
Any hospitalization	175	6.8	89	7.6
Any prescription	2,059	80.5	1,050	89.2
Conditional on use	Mean	SD	Mean	SD
Psychotherapy visits	10.8	10.3	10.8	10.9
Medication management visits	4.4	3.5	4.6	3.9
Inpatient days	14.2	12.7	16.1	17.5
Major depression	N	%	N	%
Any psychotherapy visit	5,047	48.5	3,035	57.9
Any medication management visit	3,269	31.4	1,918	36.6
Any hospitalization	290	2.8	117	2.2
Any prescription	7,590	72.9	4,319	82.4
Conditional on use	Mean	SD	Mean	SD
Psychotherapy visits	10.5	10.1	11.6	12.1
Medication management visits	3.8	3.2	3.6	2.8
Inpatient days	10.5	10.6	14.4	19.2
Adjustment disorder	N	%	N	%
Any psychotherapy visit	2,465	40.2	1,890	46.1
Any medication management visit	260	4.2	202	4.9
Any hospitalization	25	0.4	20	0.5
Any prescription	1,890	30.9	1,474	36.0
Conditional on use	Mean	SD	Mean	SD
Psychotherapy visits	11.3	9.6	10.1	10.1
Medication management visits	3.3	2.4	3.2	2.4
Inpatient days	7.2	7.5	11.6	12.7

^a FEHB Program=Federal Employees Health Benefits Program. We compared enrollees of seven FEHB Program plans with enrollees of comparison plans operated by large, self-insured employers in the Truven Health Analytics MarketScan database over the period 1999–2002.

among FEHB Program enrollees relative to the comparison group (Table 3). There was a statistically significant decrease in annual psychotherapy utilization only for enrollees in the adjustment disorder cohort (Table 4) after parity among FEHB Program enrollees relative to comparison group enrollees. There were no significant utilization changes

after parity implementation for medication management visits, prescriptions, or inpatient days. Because utilization of some services, such as inpatient bed days, was quite low, the corresponding confidence intervals are wide. Our results were unchanged when we computed the less conservative Hochberg adjustment for multiplicity.

TABLE 3. Difference-in-Difference Results (Controlling for Secular Trends) for Annual Mental Health/Substance Use Disorder Spending Outcomes Per Person for Enrollees With Bipolar Disorder, Major Depression, or Adjustment Disorder, After (2002) Compared With Before Parity Implementation (2000)^a

Spending Measure	After Parity Implementation	
	Change (\$)	99.2% CI
Bipolar disorder		
Total spending	−436	−951, 25
Out-of-pocket spending	−148	−217, −85
Major depression		
Total spending	−36	−223, 130
Out-of-pocket spending	−100	−123, −77
Adjustment disorder		
Total spending	−62	−133, −11
Out-of-pocket spending	−68	−84, −54

^a Individuals were identified for each diagnostic group based on ICD-9 diagnosis codes in 1999 claims data. The difference-in-difference results reflect changes before parity implementation (2000) compared with after parity implementation (2002) for enrollees in the Federal Employees Health Benefits Program group relative to individuals in the comparison group. Confidence intervals are adjusted for multiple comparisons equivalent to 95% CIs. Adjusted changes are not conditional on use but averaged among all enrollees in a diagnostic cohort.

Discussion

The effects on mental health/substance use disorder utilization and spending of implementing a comprehensive parity policy in the context of managed care in the FEHB Program differed across diagnoses of varying severity. While total mental health/substance use disorder spending was unchanged for enrollees with bipolar disorder and major depression after parity implementation for FEHB Program enrollees relative to comparison group enrollees, it was significantly lower on average for those diagnosed with adjustment disorder, a diagnosis considered less severe and chronic than the other two. Parity provided additional financial protection through decreased out-of-pocket costs to enrollees across all three diagnostic categories, which is consistent with previous studies of parity among all mental health/substance use disorder users in the FEHB Program and individuals subject to the Oregon state parity law (5, 6). However, the decrease in adjustment disorder out-of-pocket spending likely reflects the postparity decrease in total mental health/substance use disorder spending for this group.

Notably, there was a statistically significant decline in annual psychotherapy utilization only for individuals with adjustment disorder. While previous research has documented a secular trend in declining rates of psychotherapy (13–16), the difference-in-difference design of this study controls for secular trends. Thus, the reductions observed here represent the net effect of the FEHB Program parity policy, which included both parity in benefit design and the encouragement of benefit management.

TABLE 4. Difference-in-Difference Results (Controlling for Secular Trends) for Annual Mental Health/Substance Use Disorder Service Utilization Outcomes Per Person for Enrollees With Bipolar Disorder, Major Depression, or Adjustment Disorder, After (2002) Compared With Before Parity Implementation (2000)^a

Service Utilization Measure	After Parity Implementation	
	Change (%)	99.2% CI
Bipolar disorder		
Psychotherapy visits	−10	−24, −0.9
Medication management visits	3.8	−11, 18
Mental health/substance use disorder prescriptions	−1.7	−7.8, 4.2
Inpatient days	17	−32, 62
Major depression		
Psychotherapy visits	−4.6	−10, 0.9
Medication management visits	−1.4	−7.9, 5.3
Mental health/substance use disorder prescriptions	−0.7	−4.3, 3.0
Inpatient days	18	−39, 61
Adjustment disorder		
Psychotherapy visits	−12	−19, −4
Medication management visits	12	−26, 49
Mental health/substance use disorder prescriptions	2.7	−6.7, 12
Inpatient days	53	−196, 224

^a Individuals were identified for each diagnostic group based on ICD-9 diagnosis codes in 1999 claims data. The difference-in-difference results reflect changes before parity implementation (2000) compared with after parity implementation (2002) for enrollees in the Federal Employees Health Benefits Program group relative to individuals in the comparison group. Confidence intervals are adjusted for multiple comparisons equivalent to 95% CIs. Adjusted changes are not conditional on use but averaged among all enrollees in a diagnostic cohort.

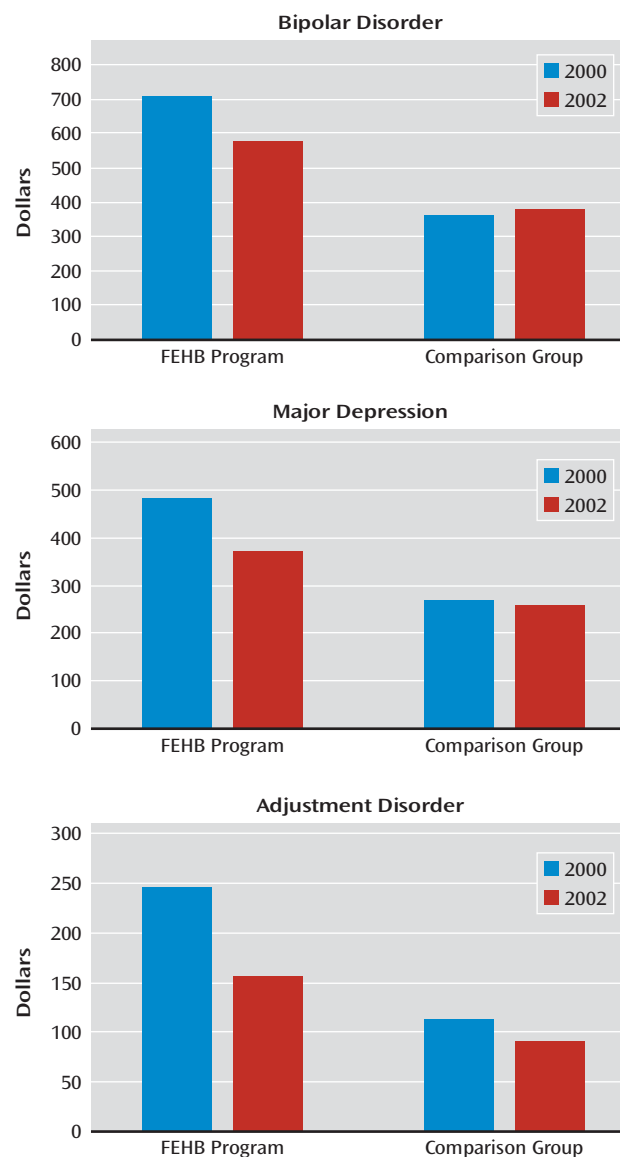
To further illustrate our findings, Figure 1 summarizes the changes in out-of-pocket spending across the three diagnostic groups by health plan, comparing preparity and postparity implementation in the FEHB Program and adjusting for enrollee characteristics (age, sex, relation to employee, and geographical region). Out-of-pocket spending attributable to parity decreased approximately 21% for each of the diagnoses. These reductions include the out-of-pocket reductions in the FEHB Program plans relative to changes in the comparison plans. Thus, for bipolar disorder, out-of-pocket savings attributable to parity are larger than the decrease seen in the FEHB Program plans (because out-of-pocket spending increased for bipolar disorder enrollees in the comparison plans), whereas for major depression, out-of-pocket spending was reduced in both FEHB Program and the comparison plans. Thus, some of the out-of-pocket savings seen in the FEHB Program for enrollees with major depression would have occurred as a result of secular trends even without parity. FEHB Program enrollees with adjustment disorder also experienced a reduction in out-of-pocket spending (as did those in the comparison plan), but given the reductions in total spending and in psychotherapy utilization for this group, much of the reduced out-of-pocket

attributable to parity reflects reduced utilization rather than increased financial protection.

In summary, we found two main statistically significant effects of the parity policy: 1) additional financial protection, particularly for those with major depression and bipolar disorder; and 2) preservation of spending and services for diagnoses that are, on average, more severe, but a reduction in spending and services for diagnoses that are expected to be less so. It is difficult to assess the financial impact of these savings for FEHB Program enrollees. The 21% out-of-pocket savings attributable to the parity policy for enrollees with bipolar disorder and major depression (the two diagnostic categories that experienced reductions in out-of-pocket spending but not in services or total spending) likely represents a meaningful cost decrease for many. FEHB Program enrollees include people with a wide range of jobs and pay grades—custodians who tend to the physical plant of federal buildings, postal workers, and administrative support staff, to name a few. However, the absolute dollar amounts (\$148 and \$100 on average, annually) are not very large, and the decrease in out-of-pocket expenses does not represent meaningful protection against catastrophic financial loss. There were few cases of extremely high out-of-pocket expenses in either period (to affect the average), and it seems likely that some of the protection against catastrophic expenses was due to the federal Mental Health Parity Act of 1996—which predates the FEHB Program parity policy and prohibited health insurers who covered mental health services from imposing higher annual or lifetime limits for these services than they did for general medical care. The FEHB Program parity policy still provided additional benefit parity, though, which is consistent with our findings of improved insurance protection from the policy.

Our utilization findings are consistent with the theory that health plans will respond to parity regulation by ratcheting up managed care to control spending increases that might otherwise accompany benefit expansion. Concerns about spending increases following benefit expansion under parity were consistent with the early research literature on this topic (17). The RAND Health Insurance Experiment, a randomized experiment of insurance benefit design on health care use and spending that was conducted in the 1970s and 1980s before managed care became widespread, found that among individuals enrolled in fee-for-service plans, decreases in enrollee cost sharing increased use of outpatient mental health services at twice the rate as for general medical outpatient services (18). In its parity directive for the FEHB Program, the OPM explicitly encouraged FEHB Program plans to use managed care techniques to control any spending increases that could result from the directive's implementation. Therefore, it is not possible in this study to separate out the effects of parity from changes in care management. We speculate, however, that our findings of decreased utilization are likely the result of an increase in benefit management that occurred

FIGURE 1. Out-of-Pocket Spending for Enrollees With Bipolar Disorder, Major Depression, or Adjustment Disorder in FEHB Program Plans and Comparison Health Plans Before (2000) and After (2002) Parity Was Implemented in the FEHB Program Plans^a



^a FEHB Program=Federal Employees Health Benefits Program. Unlike FEHB Program enrollees with bipolar disorder and major depression, those with adjustment disorder were the only group to have reduced total spending and service utilization (psychotherapy) as a result of parity. Therefore, for adjustment disorder enrollees in the FEHB Program, reduced out-of-pocket costs likely reflects reduced service use rather than improved financial protection. Out-of-pocket spending is adjusted for enrollee age, sex, dependent status, and region.

alongside implementation of parity. This is consistent with previous research showing that FEHB Program health plans increased their use of a variety of benefit management techniques after parity (e.g., primary care gatekeeping, treatment authorization, requirement of treatment plans, concurrent and retrospective reviews, and closed provider panels) (19). These techniques had been implemented in

some health plans before parity implementation to control secular spending increases. But after parity, most FEHB Program enrollees were subject to stricter benefit management than they had been subject to previously.

This finding of increased benefit management accompanying parity is highly relevant given a key difference between the FEHB Program parity directive and the MHPAEA in how plans are permitted to use managed care. Unlike the FEHB Program directive, MHPAEA regulations implemented in 2011 expressly prohibit health plans from imposing more restrictive managed care techniques for mental health/substance use disorder benefits than for other health benefits as a method of controlling spending. It is possible that this regulatory provision could affect spending in response to parity among enrollees in these three diagnostic groups.

There are several limitations to consider in this analysis. First, we determined our diagnostic cohorts based on diagnosis information on claims in the baseline year. Given the often episodic symptomatic course of these illnesses, we cannot rule out the possibility that changes in utilization postparity were due to changes in symptom course independent of parity. Similarly, in claims data we cannot observe clinical outcomes, so we are unable to determine whether the postparity policy changes in utilization were associated with differential clinical outcomes. A second consideration is that utilization patterns and care management practices have changed since the study period. For example, between 1997 and 2008, mental health/substance use disorder hospitalization rates increased 15%, while lengths of stay grew shorter by 11% (20). Third, while we did not detect changes in utilization for higher-intensity services, such as hospitalization, as a result of parity, it is likely that our study was underpowered to detect changes for this service type. Finally, our approach excluded enrollees who were not continuously enrolled all four study years, so we cannot comment on the effect of the parity policy on spending and utilization among those individuals. However, this exclusion criterion was necessary to ensure that our study results were not biased by utilization patterns reflective of changes in the enrollee population, and it allowed the use of a study design that provided more confidence that changes we observed were due to the FEHB Program parity policy rather than secular trends.

Our study provides important new information regarding the effect of implementing a comprehensive mental health/substance use disorder parity policy on diagnoses that vary in severity and chronicity. Care management likely explains why the FEHB parity policy did not translate into higher utilization across diagnoses, and our study demonstrates that benefit management was applied more stringently for illnesses that are generally expected to be less severe and chronic. This difference across diagnoses would suggest that health plan care management techniques allocated resources with at least some consideration of the clinical context or need of enrollees. The promise of managed

behavioral health care was to replace arbitrary limits and higher cost-sharing arrangements with clinically sensitive case-by-case management, including utilization review and requiring written treatment plans. However, in the absence of outcome data it is difficult to know whether postparity utilization levels were appropriate for FEHB Program enrollees. Future research must examine whether or how mental health/substance use disorder care management changes given the restrictions on management under the MHPAEA and the resulting impact on utilization and outcomes.

Presented in part at the 2011 NIMH Mental Health Services Research Conference, Washington, D.C., July 27, 2011; and at the AcademyHealth Annual Research Meeting, Orlando, Fla., June 24, 2012. Received March 26, 2012; revisions received July 18 and Oct. 1, 2012; accepted Oct. 4, 2012 (doi: 10.1176/appi.ajp.2012.12030392). From McLean Hospital, Belmont, Mass.; the Department of Health Care Policy, Harvard Medical School, Boston; Mathematica Policy Research, Washington, D.C.; Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, Baltimore; and Department of Psychiatry, University of Maryland School of Medicine, Baltimore. Address correspondence to Dr. Busch (abusch@mclean.harvard.edu).

The authors report no financial relationships with commercial interests.

The authors acknowledge funding support from NIMH (grant R01MH080797; grant K01MH071714 to Dr. Busch; grant R01MH093414 to Drs. Barry, Goldman, and Huskamp; and grant R01MH054693 to Drs. Normand and Yoon) and from the Health Services Research Division of Partners Psychiatry and Mental Health, Partners HealthCare (to Dr. Busch). The authors also thank Hocine Azeni for his programming expertise and support.

References

1. The Mental Health Parity and Addiction Equity Act. Washington, DC, US Department of Health and Human Services, Centers for Medicare and Medicaid Services, 2008 (www.cms.gov/Health-InsReformforConsume/04_TheMentalHealthParityAct.asp)
2. Barry CL, Gabel JR, Frank RG, Hawkins S, Whitmore HH, Pickreign JD: Design of mental health benefits: still unequal after all these years. *Health Aff (Millwood)* 2003; 22:127–137
3. Azrin ST, Huskamp HA, Azzone V, Goldman HH, Frank RG, Burnam MA, Normand SL, Ridgely MS, Young AS, Barry CL, Busch AB, Moran G: Impact of full mental health and substance abuse parity for children in the Federal Employees Health Benefits Program. *Pediatrics* 2007; 119:e452–e459
4. Azzone V, Frank RG, Normand SL, Burnam MA: Effect of insurance parity on substance abuse treatment. *Psychiatr Serv* 2011; 62:129–134
5. Goldman HH, Frank RG, Burnam MA, Huskamp HA, Ridgely MS, Normand SL, Young AS, Barry CL, Azzone V, Busch AB, Azrin ST, Moran G, Lichtenstein C, Blasinsky M: Behavioral health insurance parity for federal employees. *N Engl J Med* 2006; 354: 1378–1386
6. McConnell KJ, Gast SHN, Ridgely MS, Wallace N, Jacuzzi N, Rieckmann T, McFarland BH, McCarty D: Behavioral health insurance parity: does Oregon's experience presage the national experience with the Mental Health Parity and Addiction Equity Act? *Am J Psychiatry* 2012; 169:31–38
7. Zeger SL, Liang KY: Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 1986; 42:121–130
8. Buntin MB, Zaslavsky AM: Too much ado about two-part models and transformation? Comparing methods of modeling Medicare expenditures. *J Health Econ* 2004; 23: 525–542

9. Duan N: Smearing estimate: a nonparametric retransformation method. *J Am Stat Assoc* 1983; 78:605–610
10. Su L, Tom BD, Farewell VT: Bias in two-part mixed models for longitudinal semicontinuous data. *Biostatistics* 2009; 102:374–389
11. Neelon B, O'Malley AJ, Normand SL: A Bayesian two-part latent class model for longitudinal medical expenditure data: assessing the impact of mental health and substance abuse parity. *Biometrics* 2011; 67:280–289
12. Yoon FB, Fitzmaurice GM, Lipsitz SR, Horton NJ, Laird NM, Normand SL: Alternative methods for testing treatment effects on the basis of multiple outcomes: simulation and case study. *Stat Med* 2011; 30:1917–1932
13. Berndt ER, Bir A, Busch SH, Frank RG, Normand SLT: The medical treatment of depression, 1991–1996: productive inefficiency, expected outcome variations, and price indexes. *J Health Econ* 2002; 21:373–396
14. Olfson M, Marcus SC: National trends in outpatient psychotherapy. *Am J Psychiatry* 2010; 167:1456–1463
15. Busch AB, Ling D, Frank RG, Greenfield SF: Changes in the quality of care for bipolar I disorder during the 1990s. *Psychiatr Serv* 2007; 58:27–33
16. Olfson M, Marcus SC, Druss B, Elinson L, Tanielian T, Pincus HA: National trends in the outpatient treatment of depression. *JAMA* 2002; 287:203–209
17. Barry CL, Frank RG, McGuire TG: The costs of mental health care parity: still an impediment? *Health Aff (Millwood)* 2006; 25: 623–634
18. Newhouse JP; Insurance Experiment Group: *Free For All? Lessons From the RAND Health Insurance Experiment*. Cambridge, Mass, Harvard University Press, 1993
19. Ridgely MS, Burnam MA, Barry CL, Goldman HH, Hennessy KD: Health plans respond to parity: managing behavioral health care in the Federal Employees Health Benefits Program. *Milbank Q* 2006; 84:201–218
20. National Hospital Discharge Survey: Hospital discharges by first- and any-listed diagnosis: US, 1990–2007. <http://205.207.175.93/HDI/TableViewer/tableView.aspx?ReportId=537>