

Use of Performance Standards in Behavioral Health Carve-Out Contracts Among Fortune 500 Firms

Elizabeth L. Merrick, PhD, MSW; Deborah W. Garnick, ScD; Constance M. Horgan, ScD; David Goldin, MA; Dominic Hodgkin, PhD; and Mark Sciegaj, PhD

Abstract

Objective: To determine the prevalence and nature of performance standards in specialty managed behavioral healthcare contracts among Fortune 500 companies.

Study Design: This was a cross-sectional survey of all companies listed on the Fortune 500 during 1994, 1995, or both.

Methods: From April 1997 to May 1998 we conducted a mailed survey with phone follow-up. Of the 68% of firms that responded, over one third reported carve-out contracts. The survey focused on whether companies had behavioral health carve-out contracts with specialty vendors and characteristics of these contracts, including the use of performance standards.

Results: More than three quarters of the Fortune 500 companies reporting specialty behavioral healthcare contracts used at least one performance standard. Most common were administrative standards (70.2%) and customer service standards (69.4%). About half of the companies used quality standards, whereas only a third used provider-related standards. Most (58.8%) companies using performance standards also specified financial consequences. Larger Fortune 500 firms were significant-

ly more likely to use performance standards. Risk contracts and contracts that included all covered employees were also more likely to include some categories of standards.

Conclusions: Administrative and customer service standards may be most common because companies find it easier to specify those standards, especially compared with clinical quality measures. To the extent that employers want to obtain the most value from their behavioral healthcare purchasing, we expect that more will begin to adopt quality standards in their contracts, especially as performance measures become more refined. Reliance on accreditation, however, is an alternative approach for employers.

(Am J Managed Care 1999;5:SP81-SP90)

With the rise of managed care, concerns about maintaining and improving quality have come to the forefront. Measurement of health plan performance has been an increasing focus in recent years. Measurement of behavioral health performance in particular has progressed significantly, as reflected in development of consumer-oriented report cards and inclusion of behavioral health measures in health plan performance measurement systems.¹

At the same time, little is known about employers' use of performance standards in behavioral healthcare contracting. Given the proliferation of specialty behavioral healthcare and the view of performance standards as a useful tool to ensure quality and value, it is important to understand the use of these standards. This study examines the prevalence, types, and structure of performance standards used

From the Institute for Health Policy, Heller Graduate School, Brandeis University, Waltham, MA.

This research was supported by a center grant from the National Institute on Drug Abuse (grant #1 P50 DA 10233) for the Brandeis/Harvard Research Center on Managed Care and Drug Abuse Treatment.

Address correspondence to: Elizabeth L. Merrick, PhD, MSW, Institute for Health Policy, Heller Graduate School, MS 035, Brandeis University, 415 South St, Waltham, MA 02454-9110. E-mail: merrick@brandeis.edu.

by Fortune 500 firms that contracted directly with specialty behavioral healthcare vendors in 1997.

... BACKGROUND ...

Over the past decade, many changes have occurred in the ways that purchasers arrange for behavioral health services. One major development has been the move towards carve outs, in which behavioral health benefits are separated from other medical benefits and contracted for separately with a specialty vendor. Prevalence of carve outs is difficult to measure because of rapid changes, varying definitions, and double counting of enrollees in multiple types of managed behavioral healthcare. However, one recent estimate is that in 1997, 149 million persons were enrolled in specialty behavioral health programs, excluding internally managed health maintenance organization (HMO) services.²

Several have noted key issues in the rationale for behavioral health carve outs.^{3,4} First, use of carve outs may lead to lower costs for mental health and substance abuse services for a variety of reasons, including tighter management and shifting to lower-cost providers. Second, the specialty expertise of carve-out vendors also may facilitate high quality of care. Third, purchasers are able to allocate resources specifically to behavioral health coverage, which has been limited in part by health plans' concern about adverse selection.

Contracting with specialty managed behavioral healthcare vendors in carve-out arrangements can occur at 2 levels.⁵ Purchasers such as employers or Medicaid programs may contract directly for behavioral healthcare. Among employers, large employers are especially likely to carve out these services.⁶ Health plans such as health maintenance organizations also may choose to contract out the management and delivery of behavioral health services in carve-out arrangements.

Another innovation has been the use of performance standards in behavioral healthcare contracts. Performance standards identify acceptable levels of performance for various aspects of health service delivery, including both administrative and clinical responsibilities. For example, vendors might be required to maintain a specified ratio of mental health specialists to enrollees, or to answer 95% of member service phone calls within 30 seconds. Financial penalties for failure to meet a standard, rewards for exceeding the standard, or both may be

included. The use of performance standards in contracting can be seen as a reflection of more sophisticated healthcare purchasing by some employers and other purchasers. Performance standards formalize purchaser expectations and vendor accountability. Specifying standards requires that the purchaser have clear ideas as to what is important and feasible to monitor.

Purchasers wishing to specify performance standards that address clinical quality can create their own or borrow those developed by others. Accreditation organizations such as the National Committee for Quality Assurance (NCQA) have developed behavioral health performance measures.⁷ One of the most prominent performance measurement systems is NCQA's Health Plan Employer Data and Information Set (HEDIS), which includes some behavioral health measures in the current version. Also, NCQA recently began to offer accreditation specifically for specialty managed behavioral healthcare organizations. The Foundation for Accountability also has developed a set of health plan performance measures that now includes behavioral health measures.⁸

The American Managed Behavioral Healthcare Association has developed Performance-Based Measures for Managed Behavioral Healthcare Programs, a performance measure set for behavioral health only.⁹ Some companies and government agencies, including Digital Equipment Corporation and the Substance Abuse and Mental Health Services Administration, similarly have developed behavioral health plan report cards or measurement sets.¹ Most recently, the Washington Circle Group (national experts in substance abuse, managed care, and performance management) has begun to develop a core set of substance abuse performance measures.¹⁰ All of these efforts are geared towards making better information available to consumers and institutional purchasers and to enable more informed decision making. Ultimately, the goal is to encourage health plans to achieve a high level of quality.

Despite the increasing focus on performance measurement, some evidence suggests that employers tend to consider certain quality information such as clinical outcomes, NCQA accreditation, and HEDIS measures less important than other factors such as cost or claims processing in selecting general medical plans.¹² On the other hand, employee satisfaction and number and quality of physicians ranked high in importance. In a study on how large employers use performance information in general

health plan purchasing decisions, Hibbard and colleagues found that many employers were unaware of available clinical quality measures.¹² Employers also considered some measures flawed or irrelevant and furthermore had difficulty integrating numerous different measures into their decision making.

One survey that examined the use of performance standards in employers' contracts with general health plans found that fewer than half included any performance standards.¹³ Administrative and customer service standards related to claims processing and phone response were most prevalent, whereas HEDIS measures and provider-related standards were uncommon. Another survey, by the Washington Business Group on Health and Watson Wyatt Worldwide, investigated the types of information employers used to evaluate and select healthcare plans.¹⁴ Employer-specified performance standards or goals were used by 42% of all firms and by 74% of large firms. The content of these standards was not indicated. However, cost (95% of all firms), member services to employees (84%), employee access (80%), and customer service to employers (75%) were the most commonly reported types of information used in evaluating health plans. Accreditation status was used by 49%, consumer satisfaction survey results by 41%, quality improvement initiatives by 35%, HEDIS by 28%, and disease-specific outcomes by 19% of all firms. Large firms used these types of information more often than medium-sized or small firms.

There is very little systematic information on employers' use of performance standards in specialty behavioral healthcare contracts. Published evidence appears limited to one survey report, which described typical standards but did not provide summary statistics,¹⁵ as well as case examples and general description such as that by Feldman.¹⁶ As part of a survey of Fortune 500 firms' use of behavioral health carve outs, we gathered information on performance standards. This analysis is designed to increase our knowledge of how large employers use performance standards when directly contracting for specialty behavioral healthcare.

... METHODS ...

From April 1997 to May 1998, we surveyed the 537 firms listed among the Fortune 500 during 1994 or 1995. Some firms were ineligible due to mergers or acquisitions, resulting in a total of 496 eligible

firms. The survey was mailed to all companies, and respondents had the option of completing the survey in writing or over the phone. Extensive phone follow-up was used to encourage firms to respond and to clarify some answers. Sixty-eight percent of eligible firms ($n = 336$) responded; this is one of the highest reported response rates for surveys of large employers regarding healthcare benefits.

A total of 132 firms (39% of all respondents) reported contracting with behavioral healthcare vendors in 1997. Contracts solely for employee assistance program (EAP) services were not included in this count. Of these firms, 124 responded to at least some of the questions regarding performance standards and are therefore included in this analysis.

The survey asked about various features of firms' carve-out contracts, including level of vendor risk, payment arrangements, use of performance standards, and use of financial consequences connected to performance standards. We defined vendor risk as risk for claims costs and for this analysis combined full- and partial-risk categories because so few contracts placed the vendor at full risk. Payment arrangements were classified as either administrative services only (ASO)—in which the vendor was paid for management but not treatment delivery (except EAP services)—or a combined administrative and treatment delivery payment. Some ASO contracts did involve risk for claims costs, however.

We asked firms whether their contracts with vendors specified each of 12 types of performance standards and then grouped the responses into 4 categories:

- Administrative standards included claims processing and administrative reporting.
- Customer service standards included member service phone response, complaints, appeals, and referral speed.
- Quality standards included quality assurance system standards, HEDIS behavioral health measures, enrollee satisfaction surveys, and other quality standards (including Performance-Based Measures for Managed Behavioral Healthcare Programs).
- Provider-related standards included provider satisfaction surveys and facility/staffing requirements.

Respondents were asked to report any other standards as well, which we coded into the above categories. We also asked whether financial consequences were specified in connection with performance standards.

Several variables were created by using secondary data. Revenue rank of each firm was calculated based on 1996 revenues reported on the Fortune 500 web site¹⁷ or on Hoover's web site¹⁸ for those firms no longer listed in the Fortune 500 by 1996. Firms were assigned to industry categories based on Standard Industrial Classification codes, using information from Dun and Bradstreet. Location of the firm headquarters was used to assign firms to 1 of the 4 census regions.

We present descriptive statistics for prevalence, types, and structure of performance standards. Bivariate tests (chi-square) were used to test for differences in performance standard use by firm and contract characteristics. Characteristics of our sample are shown in Table 1.

... RESULTS ...

Prevalence of Performance Standards

Table 2 shows the prevalence of performance standard use among respondent firms with carve-out contracts. More than three quarters (76.6%) of these firms included one or more performance standards in the contract on which they reported, and just over half (54.5%) included standards in the quality category. More commonly used were administrative (70.2%) and customer service (69.4%) standards, whereas provider-related standards were least common (34.7%). Table 2 also shows the prevalence of performance standard categories as the percentage of those firms that used at least one standard.

Within each category, there was variation in the prevalence of specific standards. In the administrative category, the most frequently used standard was claims processing (58.9% of all contracts). Phone response (65.3%) was the most common standard within the customer service category. In the provider-related category, provider satisfaction survey standards were most frequently reported (23.4%), whereas among the quality standards, enrollee satisfaction was most prevalent (42.7%). Overall, the 5 most frequently used standards were phone response, claims processing, administrative reporting, enrollee satisfaction, and complaints. When firms reported using any performance standards, they tended to use several (mean = 5.6, median = 5; data not shown).

Specification of Financial Consequences

The majority (59%) of the carve-out firms using performance standards specified financial consequences related to these standards. As shown in the Figure, 20 firms (21%) indicated penalties, nearly all in the form of reductions in the administrative fee if standards were not met. Nineteen (20%) noted that some portion of the administrative fee was "at risk" without explicitly stating whether rewards were possible in addition to penalties. Only 3 firms (3%) mentioned rewards in addition to penalties. Fourteen (15%) included financial consequences but did not indicate what type.

We also analyzed the relationships between the use of financial consequences and various firm and contract characteristics. In contracts with performance standards, the top 2 vendors were more likely to include financial consequences (vendor 1 = 74.3%, vendor 2 = 69.2%, other vendors = 44.7%; $P < 0.05$ [data not shown]). There was

Table 1. Description of Fortune 500 Firms Reporting Carve-Out Contracts*

Firm Characteristics	No. (%) of Responding Firms
Region	
Northeast	44 (35.5)
Midwest	31 (25.0)
South	32 (25.8)
West	17 (13.7)
Total	124 (100.0)
Industry	
Manufacturing	51 (41.1)
Transportation/communications/utilities	32 (25.8)
Wholesale/retail trades	11 (8.9)
Financial/insurance/real estate	16 (12.9)
Other industries	14 (11.3)
Total	124 (100.0)
No. of Covered Employees	
<10,000	33 (28.2)
10,000-19,999	35 (29.9)
20,000-49,999	30 (25.6)
50,000+	19 (16.2)
Total	117 (99.9)

*Note: The number of firms varies by category because of missing information. Individual categories may not total 100% because of rounding.

no significant association with revenue rank, number of covered employees, industry region, vendor risk, ASO basis, or inclusion of all covered employees.

Association of Performance Standard Use and Firm Characteristics

Table 3 presents findings on use of performance standard categories according to selected firm characteristics. Performance standards were more prevalent in firms with higher revenue and more US-based employees covered by health insurance. However, neither the headquarters location nor the length of time that companies had carved out behavioral health (data not shown) was associated with the use of performance standards.

The top 100 firms by revenue were significantly more likely to include at least one performance standard (88.6% vs 71.9%; $P < 0.05$). They were also more likely to use standards in the administrative (82.3% vs 65.2%; $P < 0.10$), customer service (88.2% vs 61.8%; $P < 0.01$), and quality (79.4% vs 43.2%; $P < 0.01$) categories. We found similar results when firm size was measured by number of covered employees. Firms with above the median number of covered employees were more likely to use at least one performance standard (84.5% vs 66.1%; $P < 0.05$). They were also more likely to use administrative (79.3% vs 59.3%; $P < 0.05$), customer service (81.0% vs 55.9%; $P < 0.01$), and quality (63.2% vs 44.1%; $P < 0.05$) standards.

Use of quality standards differed significantly ($P < 0.05$) by industry category, ranging from 25% of financial/insurance/real estate firms to 71.4% in the "other industries" category. There were no significant dif-

Table 2. Performance Standards in Fortune 500 Firms' Carve-Out Contracts

Type of Standard	Percentage of All Carve-Out Contracts (n = 124)	Percentage of Carve-Out Contracts Using Any Standards (n = 95)
Any	76.6	100.0
Administrative	70.2	91.6
Claims processing	58.9	76.8
Administrative reporting	53.2	69.5
Customer Service	69.4	90.5
Phone response	65.3	85.3
Complaints	37.9	49.5
Appeals	37.1	48.4
Referral speed	34.7	45.3
Provider Related	34.7	45.3
Provider satisfaction survey	23.4	30.5
Facilities/staffing	21.0	27.4
Quality	54.5	71.3
Enrollee satisfaction	42.7	55.8
Quality assurance	35.5	46.3
HEDIS behavioral health measures	14.6	18.9
Other	6.5	8.4

Figure. Percentage of Contracts with Performance Standards, Using Various Financial Consequences (n = 95)

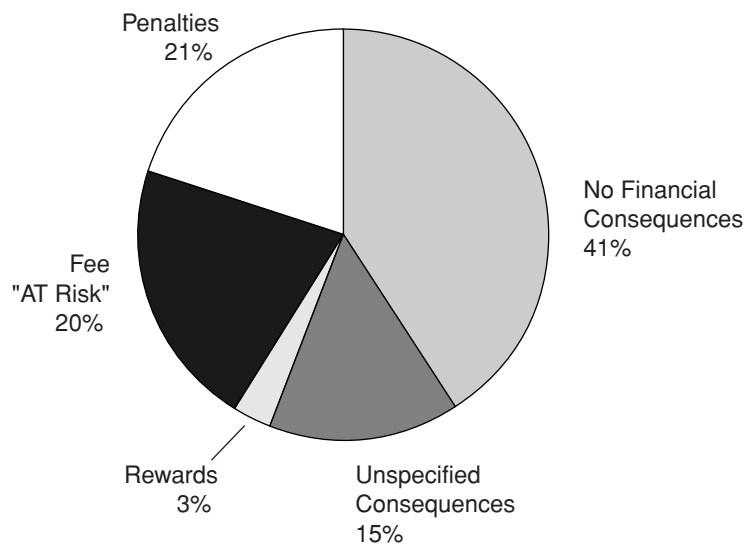


Table 3. Percentage of Fortune 500 Firms Using Categories of Performance Standards Among Those Reporting 1997 Specialty Behavioral Health Contracts

Firm Characteristics	No. of Firms*	Performance Standard Categories (%)				
		Any	Administrative	Customer Service	Quality	Provider Related
Total	124	76.6	70.2	69.4	54.5	34.7
Revenue Rank						
Top 100	35	88.6 [†]	82.3 [†]	88.2 [§]	79.4 [§]	44.1
Others	89	71.9	65.2	61.8	43.2	29.2
No. of Covered Workers						
Above median	59	84.5 [†]	79.3 [†]	81.0 [§]	63.2 [†]	37.9
Below median	58	66.1	59.3	55.9	44.1	30.5
Industry						
Manufacturing	51	76.5	71.4	85.7	52.0 [†]	33.3
Transportation/communications/ utilities	32	81.3	68.6	72.5	68.8	46.9
Wholesale/retail trades	11	72.7	78.1	71.9	45.5	9.1
Financial/insurance/real estate	16	62.5	72.7	63.6	25.0	25.0
Other	14	85.7	56.3	43.7	71.4	42.9
Region						
Northeast	44	75.0	65.9	70.5	48.8	29.6
Midwest	31	83.9	80.0	74.2	61.3	35.5
South	32	75.0	65.6	65.6	56.3	43.7
West	17	70.6	70.6	64.7	52.9	29.4

*Shown are the number of firms for which relevant data were available.

Using a chi-square test, differences among categories of revenue rank, number of covered workers, industry, or region are significant at:
[†]P < 0.05.

[†]P < 0.10.

[§]P < 0.01.

Table 4. Percentage of Fortune 500 Firms Using the Top 5 Performance Standards Among Those Reporting 1997 Specialty Behavioral Health Contracts, by Firm Characteristics

Firm Characteristics	No. of Firms*	Top 5 Performance Standards (%)				
		Phone Response	Claims Processing	Administrative Reporting	Enrollee Satisfaction	Complaints
Total	124	65.3	58.9	53.2	42.7	37.9
Revenue Rank						
Top 100	35	78.8 [†]	66.7	63.6	60.6 [†]	36.4
Others	89	60.0	57.8	47.8	35.6	37.8
No. of Covered Workers						
Above median	59	81.4 [§]	62.7	64.4 [†]	52.4 [†]	37.3
Below median	58	51.7	51.7	44.8	34.5	37.9
Industry						
Manufacturing	51	66.7 [†]	64.7	47.1	41.2 [†]	33.3 [†]
Transportation/communications/ utilities	32	71.9	68.7	68.8	56.3	56.3
Wholesale/retail trades	11	63.6	36.4	54.5	27.3	27.3
Financial/insurance/real estate	16	31.3	37.5	43.8	18.8	12.5
Other	14	85.7	57.1	50.0	57.1	50.0
Region						
Northeast	44	63.6	52.3	50.0	38.6	38.6
Midwest	31	67.7	67.7	58.1	41.9	35.5
South	32	65.6	62.5	50.0	50.0	40.6
West	17	64.7	52.9	58.8	41.2	35.3

*Shown are the number of firms for which relevant data were available.

Using a chi-square test, differences among categories of revenue rank, number of covered workers, industry, or region are significant at:

[†]P < 0.10.

[†]P < 0.05.

[§]P < 0.01.

ferences by industry in the proportion of firms using any performance standards. Geographical region of firm headquarters was not associated with differences in the use of performance standards.

Table 4 displays the breakdown of use of the top 5 specific performance standards by firm characteristics. Firms in the top 100 revenue group were significantly more likely to use phone response (78.8% vs 60.0%; $P < 0.10$) and enrollee satisfaction (60.6% vs 35.6%; $P < 0.05$) standards. Firms with an above-median number of covered employees were more likely to use phone response (81.4% vs 51.7%; $P < 0.01$), administrative reporting (64.4% vs 44.8%; $P < 0.05$), and enrollee satisfaction (52.4% vs 34.5%; $P < 0.05$) standards. Industry type was significantly related to use of enrollee satisfaction ($P < 0.10$) and complaints ($P < 0.05$), and phone response ($P < 0.05$) standards. Region was not significantly associated with any of the top 5 standards.

Association of Performance Standard Use and Contract Features

Table 5 shows the percentages of firms using performance standard categories according to selected

contract features. In contracts that placed the vendor at some risk for claims costs, administrative standards were more likely (82.8% vs 65.2%; $P < 0.10$). Risk contracts were also more likely to contain provider-related standards (51.7% vs 29.3%; $P < 0.05$). There was no difference in the use of any standards. ASO contracts were less likely than those based on payment for both administrative and treatment services to include provider-related standards (30.8% vs 56.3%; $P < 0.05$), but no other differences were evident. Contracts that included all (versus some) covered employees were significantly more likely to include administrative standards (84.0% vs 65.6%; $P < 0.10$). The specific vendor variable was associated with provider-related standards use (vendor 1 = 21.4%, vendor 2 = 35.3%, other vendors = 42.9%; $P < 0.10$).

Table 6 shows the relationship between contract features and use of the top 5 individual standards. Risk contracts were significantly more likely to contain administrative reporting standards (69.0% vs 46.7%; $P < 0.05$). ASO contracts were less likely to contain these standards (49.5% vs 75.0%; $P < 0.10$). Contracts that included all covered employees were

Table 5. Percentage of Fortune 500 Firms Using Categories of Performance Standards Among Those Reporting 1997 Specialty Behavioral Health Contracts, by Contract Features

Contract Features	No. of Firms*	Performance Standard Categories (%)				
		Any	Administrative	Customer Service	Quality	Provider Related
Total	124	76.6	70.2	69.4	54.5	34.7
Vendor Risk						
Full/partial	29	85.7	82.8 [†]	72.4	62.1	51.7 [†]
None	92	74.0	65.2	67.4	50.5	29.3
Payment Arrangement						
Administrative services only	107	75.7	68.2	70.1	52.8	30.8 [†]
Treatment delivery and administrative services	16	81.3	81.3	62.5	62.5	56.3
Employees Included						
All	25	80.1	84.0 [†]	76.0	66.7	44.0
Some	96	75.0	65.6	66.7	50.0	32.3
Specific Vendor						
Vendor 1	42	80.9	78.6	73.8	54.8	21.4 [†]
Vendor 2	17	76.5	70.6	76.5	58.8	35.3
Others	63	73.0	63.5	63.5	51.6	42.9

*Shown are the number of firms for which relevant data were available.

Using a chi-square test, differences among categories of vendor risk, payment arrangement, employees included, and specific vendor are significant at:

[†] $P < 0.10$.

[‡] $P < 0.05$.

much more likely to contain administrative reporting standards (80.0% vs 45.8%; $P < 0.01$). Contracts with the top 2 vendors were significantly more likely to include claims processing standards (vendor 1 = 71.4%, vendor 2 = 70.6%, other vendors = 47.6%; $P < 0.05$).

... DISCUSSION ...

Use of performance standards in behavioral health carve-out contracts among Fortune 500 companies is common, although not universal. The largest firms were even more likely to use performance standards, further underlining the role that size of company might play. Larger firms may tend to have greater expertise in healthcare purchasing, either in-house or through the use of consultants, that would lead them to specify performance expectations. These companies also may have more resources with which to monitor standards.

Administrative and customer service standards were more prevalent than quality measurement or

provider-related standards. It may be that administrative and customer service performance impact most immediately on firms in terms of contract management or employee complaints. Administrative and customer service requirements are also relatively simple to specify. Specifying clinical standards is more difficult because of uncertainties about the best processes of care and the influence on outcomes of multiple factors outside the control of behavioral healthcare providers. Thus, in writing contracts with behavioral healthcare firms, employers face both clear reasons to include performance standards in order to ensure good quality of care and uncertainty about whether current measurement tools are adequate.

Administrative and customer service tasks are important to perform well, so the high prevalence of standards in these areas seems reasonable and desirable. But one of the main questions raised by the study is what it means for consumers when half of the reported carve-out contracts in our sample did not include what we termed quality standards. It may be that as behavioral health measures become more refined and also more commonplace, they will

Table 6. Percentage of Fortune 500 Firms Using the Top 5 Performance Standards Among Those Reporting 1997 Specialty Behavioral Health Contracts, by Contract Features

Contract Features	Top 5 Performance Standards (%)					
	No. of Firms*	Phone Response	Claims Processing	Administrative Reporting	Enrollee Satisfaction	Complaints
Total	124	65.3	58.9	53.2	42.7	37.9
Vendor Risk						
Full/partial	29	72.4	69.0	69.0 [†]	44.8	44.8
None	92	62.0	54.3	46.7	41.3	34.8
Payment Arrangement						
Administrative services only	107	65.4	57.9	49.5 [‡]	42.1	37.4
Treatment delivery and administrative services	16	62.5	62.5	75.0	43.7	37.5
Employees Included						
All	25	76.0	68.0	80.0 [§]	52.0	44.0
Some	96	62.5	55.2	45.8	38.5	36.5
Specific Vendor						
Vendor 1	42	71.4	71.4 [†]	59.4	42.9	31.0
Vendor 2	17	76.5	70.6	58.8	41.2	47.1
Others	63	57.1	47.6	46.0	41.3	38.1

*Shown are the number of firms for which relevant data were available.

Using a chi-square test, differences among categories of vendor risk, payment arrangement, employees included, and specific vendor are significant at:

[†] $P < 0.05$.

[‡] $P < 0.10$.

[§] $P < 0.01$.

more frequently appear as contractually based standards. Also, it may be more difficult to include quality standards because the data for the measures are not yet routinely collected. Performance measures may, furthermore, branch out into forms that employers would find especially relevant—for instance, translating clinical outcomes into productivity terms.

Alternatively, purchasers can establish accountability for quality of care in a variety of ways. For example, they could require detailed reports on certain indicators (eg, access, utilization, readmission rates) without specifying standards concerning those areas. The information could in fact be contained in the administrative reporting standard that we found to be common. Purchasers could then use those data to evaluate aspects of service delivery, give feedback to the vendor regarding any concerns, and factor the data into the decision about contract renewal or continuation. Research is needed to evaluate the relative impact of various approaches to signaling purchaser expectations regarding quality standards.

Firms also might view the specification of quality standards as micromanagement of care processes that are purposely being delegated in the first place. Purchasers may believe that they have chosen high-quality vendors and do not need to specify quality expectations in a detailed fashion. Of course, even a comprehensive set of quality indicators can only address a small portion of the clinical situations that arise in practice. Indeed, as Eddy points out, plans tend to allocate more resources to activities for which performance is being monitored.¹⁹ He suggests rotation of measures as one way to ensure that ongoing “distortion” does not occur. Certainly, accurate choice of the most important activities to measure is crucial; otherwise, the standards may drive inefficient resource allocation without necessarily improving overall quality. At the same time, lack of any performance standards might reflect a tendency toward less than adequate oversight of delegated responsibilities.

A substantial minority of firms using performance standards did not specify any financial consequences. This suggests that firms may view standards as primarily a feedback mechanism, providing information that can then be used in the reassessment process of whether to continue contracting with a particular vendor. It may be that the incentive of continued business (and the threat of lost business) is seen as more powerful than relatively

small financial consequences attached to individual standards. This is reminiscent of Ma and McGuire’s hypothesis of a “reputation effect” on the vendor side to explain why a carve-out vendor far exceeded the explicit financial incentives in a risk-sharing contract.²⁰

Interestingly, risk-based contracts were more likely to contain administrative and provider-related standards. Possibly, employers perceive more of a need to monitor performance under risk contracting because they are aware that such incentives may lead to skimping on aspects of service management or delivery. However, despite the possible economic incentives to undertreat, risk contracts did not contain significantly more quality standards than non-risk contracts. When we explored the use of performance standards using 3 categories of vendor risk (full, partial, none), results were generally similar. The finding that ASO contracts are less likely to contain provider-related standards is consistent with the finding regarding risk. Although some ASO contracts involve risk, many do not.

We also found that contracts with the top 2 vendors were less likely to include provider-related standards and more likely to attach financial consequences to standards. It is unknown whether this pattern stems more from the firm side (with firms having different preferences for contract specifications depending on the vendor) or from the vendor side (in which certain standards are routinely included). Future work will report on the distribution and “clustering” of particular standards within vendors in order to investigate this issue. Large vendors may be in a better financial position than smaller vendors to offer performance guarantees by taking on financial risk for whatever standards are negotiated.

Certain limitations of our study should be noted when interpreting our data. The results of surveying the largest companies in the United States cannot be generalized to smaller employers, although the direction of the differences can be plausibly hypothesized. We did not include contracts solely for EAP services, which are also part of the specialty behavioral health arena. Also, for firms with more than one carve-out contract, we obtained information on the largest contract only; however, very few reported more than one contract.

... CONCLUSIONS ...

This study found that use of performance standards was common among Fortune 500 firms

reporting a 1997 behavioral health carve out, with about three quarters of firms specifying at least one standard. These firms most often reported administrative and customer service standards; quality and provider-related standards were less often described.

Performance standards in behavioral healthcare are still at a relatively early stage, particularly in terms of clinical quality measures. To the extent that employers want to obtain the most value from their healthcare purchasing, we expect that more will begin to adopt quality standards in their contracts. This should be especially true as behavioral health measures become more refined. Another possible scenario, however, is that accreditation specifically for specialty managed behavioral healthcare organizations (such as that recently begun by NCQA) will help to establish a standard, publicly known "floor"—or minimum level—of quality. Employers may come to perceive this accreditation status as an assurance of quality that obviates the need for explicit, contractual performance standards. Finally, we recognize that the literature on general healthcare purchasing suggests that measurable quality is not always, or even typically, the determining factor. If this holds true for behavioral health, then we might expect much more gradual change in the way employers approach the quality issue.

Acknowledgments

We appreciate the participation of the Fortune 500 benefits managers and other personnel who responded to our survey. We would like to thank Isaac Montoya, Claire Sharda, Laura Altman, Suzanne Gelber, Richard Frank, Bruce Davidson, Haiden Huskamp, Matthew Urato, Eric Steinert, Anne Brisson, Dennis McCarty, and our advisory committee for helpful comments and assistance with the survey. We also thank Michele King, Lisa Andersen, Eliot Banford, and Terri White for providing administrative support.

... REFERENCES ...

1. Institute of Medicine. *Managing Managed Care: Quality Improvement in Behavioral Health*. Washington, DC: National Academy Press; 1997.
2. Oss ME, Drissel AB, Clary J. *Managed Behavioral Health Market Share in the United States, 1997-1998*. Gettysburg, PA: Open Minds; 1997.
3. Frank RG, McGuire TG, Newhouse JP. Risk contracts in managed mental health care. *Health Aff* 1995;14(3):50-64.
4. Horgan C. Financing and Reimbursement Arrangements for Provision of Alcohol Services. Paper prepared for NIAA Advisory Council. December 1995.
5. Hodgkin D, Horgan C, Garnick D. Make or buy: HMOs contracting arrangements for mental health care. *Adm Policy Ment Health* 1997;24(4):359-376.
6. Umland B. Behavioral healthcare benefit strategies of self-insured employers. *Behav Healthcare Tomorrow* 1995;6(6):67-70.
7. National Committee for Quality Assurance. Behavioral health performance measures. Available at: <http://www.ncqa.org/pages/communications/news/h99meas.htm> (HEDIS 1999 Reporting Set Measures). Accessed: Jan. 12, 1999.
8. Foundation for Accountability. Measuring quality. Available at: <http://www.facct.org/measures.html>. (Measuring Quality). Accessed: Feb. 2, 1999.
9. American Managed Behavioral Healthcare Association. Performance-Based Measures for Managed Behavioral Healthcare Programs. Available at: <http://www.ambha.org.ambhaacc.htm>. (AMBHA accomplishments). Date Accessed: Feb. 1, 1999.
10. Garnick DW, McCorry, F, Chalk M, Cotter F. *Improving Performance Measurement by Managed Care Plans for Substance Abuse: Year 1 Report of the Washington Circle Group*. Executive Summary. Rockville, MD: Center for Substance Abuse Treatment, March 16, 1999.
11. Gabel JR, Hunt KA, Hurst K. *When Employers Choose Health Plans: Do NCQA Accreditation and HEDIS Data Count?* New York, NY: KPMG Peat Marwick, LLP; September 1997.
12. Hibbard JH, Jewett JJ, Legnini MW, et al. Choosing a health plan: Do large employers use the data? *Health Aff* 1997;16(6):172-180.
13. Deloitte, Touche LLP. 1997 Employer survey on managed care. *Medical Benefits* 1998;15(10):1-2.
14. Washington Business Group on Health, Watson Wyatt Worldwide. *Partnership or One-Time Purchase, Third Annual Survey Report on Purchasing Value in Health Care*. Bethesda, MD: Watson Wyatt Worldwide, 1998.
15. Oss ME. What performance standards are used to evaluate managed behavioral health plans? *Open Minds* March 1994;7(12):4-5.
16. Feldman S. Behavioral health services: Carved out and managed. *Am J Managed Care* 1998;4(special issue):SP59-SP67.
17. The Fortune 500: our annual list of America's largest companies. *Fortune* 1997;135(8):F1.
18. Hoover's, Inc. <http://www.hoovers.com/search.html>. Accessed: Jan. 27, 1998.
19. Eddy DM. Performance measurement: Problems and solutions. *Health Aff* 1998;17(4):7-25.
20. Ma CA, McGuire TG. Costs and incentives in a behavioral health carve-out. *Health Aff* 1998;17(2):53-69.