MEDICAL CARE
Volume 37, Number 2, pp 180–188
©1999 Lippincott Williams & Wilkins, Inc.

Assessing Quality of Care

Administrative Indicators and Clinical Outcomes in Posttraumatic Stress Disorder

ROBERT ROSENHECK, MD,*†‡ ALAN FONTANA, PHD,*† AND MARILYN STOLAR, MPHIL*‡

BACKGROUND. Although the use of quality of care indicators based on data collected for administrative purposes has become widespread, the relationship between those measures and clinical outcomes has yet to be evaluated.

RESEARCH DESIGN. This study used hierarchical linear modeling to examine the relationship between 12 performance indicators derived from administrative data sets and 6 clinical outcome measures addressing symptoms, substance abuse, and social functions.

SUBJECTS. Patient interviews were conducted with 4,165 veterans 4 months after their discharge from 62 specialized VA inpatient programs for treatment of Posttraumatic Stress disorder.

RESULTS. Five of twelve administrative measures were significantly associated with at least one of the clinical outcome measures, which was all in the expected directions. The number of hospital readmissions during the 6 months after the index discharge was significantly related to poor outcomes on all 5 of 6 measures. Measures of readmission and post-discharge hospital use were more strongly and consistently related to outcome than to measures of access, intensity, or continuity of outpatient care.

CONCLUSION. Administrative data, especially measures of hospital readmission, are significantly related to clinical outcomes. Correlations, however, are small to modest in magnitude indicating that these 2 types of performance measures assess different aspects of quality and can not be substituted for one another.

Key words: mental health; quality of care; outcome assessment; PTSD; Department of Veterans Affairs. (Med Care 1999;37:180–188)

In recent years, as management procedures for controlling health service utilization and reducing health care costs have come into widespread use, concerns have intensified so much so that the quality and effectiveness of health care services may suffer.[1–6] In response to those concerns, the development of methods for monitoring health care quality and outcome has become a high priority for health systems managers and services researchers.[1]

Although methods to evaluate the efficacy of psychiatric and substance abuse treatment are well developed, the application of those methods to performance evaluation of general psychiatric practice is still in its infancy.[7] Crucial differences between the demands of efficacy research and of health system performance monitoring are as follows: (1) performance data must address outcomes for much larger numbers of patients; (2) the units of comparison in performance monitoring

*From the Northeast Program Evaluation Center, VAMC West Haven, Connecticut.

†From the Yale University Department of Psychiatry, VAMC West Haven, Connecticut.

‡From the Yale University School of Epidemiology and Public Health, West Haven, Connecticut.

Address correspondence to: Robert Rosenheck MD/182, Northeast Program Evaluation Center, 950 Campbell Avenue, West Haven, Connecticut 06516.

Received April 28, 1998; initial review completed June 19, 1998; final acceptance July 28, 1998.

are tens to hundreds of health care institutions or providers rather than a few treatment groups; and (3) random assignment to treatment condition is impossible. In a treatment efficacy study, for example, data on the impact of an intervention on several hundred randomly assigned patients are rigorously evaluated and generalized to the entire adult population without time specification. In outcomes monitoring, in contrast, treatment provided by numerous providers to hundreds or thousands of patients is evaluated and the temporal frame for generalization may be no more than 1 or 2 years. Thus, whereas expenditures of $5,000 to $10,000 for each patient that was studied are not unreasonable in the context of treatment efficacy research, such costs are totally impractical in the realm of performance assessment.

In view of the demand for large quantities of time-specific and provider-specific performance monitoring data, many health systems rely on inexpensive outcome measures derived from administrative or claims data,[8] or one-time satisfaction surveys.[9-10] The Health Employer Data Information Set (HEDIS), the basis for most evaluation of managed care networks, relies almost entirely on indicators derived from administrative data (eg, readmission rates, total annual hospital days per capita, and time from hospital discharge to first outpatient visit).[8] Such data are inexpensive to obtai,n and data collection can be well standardized across providers.

These measures, however, have not yet been validated with respect to clinical outcomes. A recent compendium of outcome measures of potential value in health system performance assessment made minimal mention of this type of measure[7] and focused, instead, on established research measures. The contrast between the inexpensive measures that receive the widest use and the more sophisticated measures that demonstrate reliability and validity is striking. It is, thus, not yet known whether less expensive indicators based on administrative data can be used as proxies for more established clinical outcome measures. We hypothesized that as administrative measures are selected to reflect quality of care and as quality of care is presumed to influence health status improvement as detected through outcome assessments, administrative measures would be associated with clinical improvement. In this study we test that hypothesis by examining the correspondence of these two types of measures to each other.

In 1995 the Department of Veterans Affairs implemented a multi-component National Mental Health Program Performance Monitoring System to evaluate the care provided to over 500,000 patients per year receiving specialized psychiatric care at over 150 medical centers across the country.[11-12] That system is based on administrative data of the HEDIS type and an annual system-wide patient satisfaction survey[13] and has been used to guide and document steady improvements in VA mental health care from 1994 to 1997.[12] A recent comparison of VA performance with that of private sector care documented in the Marketscan Database(R) of the Medstat Group (Washington, DC) for the same years, has shown VA care to be equal or better than private sector care on most measures (manuscript in preparation). A special subcomponent of the system, however, collects outcome interview data at the time of program entry and 4 months after patients were discharged among patients admitted to specialized inpatient units that treat war-related Posttraumatic Stress Disorder (PTSD), a condition of special importance in the VA.

In this study, we examine the relationship between performance measures derived from administrative data sets and outcome measures based on patient questionnaires conducted before and after treatment in a sample of 4,165 veterans discharged from 62 specialized inpatient PTSD programs between May, 1993, and March, 1996. In this study, we address the following three questions: (1) Are individual performance indicators derived from administrative data significantly correlated with any of several outcome measures obtained through patients' self-reports before and after treatment?; (2) Which specific administrative indicators are most strongly associated with clinical outcomes?; and (3) Can inexpensive administrative indicators be used as proxy measures of clinical outcome?

Methods

Performance Assessment of VA Inpatient Programs Using Administrative Data

As part of the VA's National Mental Health Program Performance Monitoring system[11-12] administrative data were used to monitor the quality of care provided to patients discharged from VA psychiatric inpatient programs. Twelve measures,

addressing both inpatient care and post-discharge outpatient follow-up care, were determined to be appropriate to the evaluation of the performance of specialized inpatient and residential PTSD programs.[14] The data for those monitors were obtained from 2 VA administrative files: the Patient Treatment File, which was a discharge abstract file on all completed episodes of VA inpatient care, and the Outpatient Care file, which was a record of outpatient services provided by VA clinics.

Outcomes Monitoring of VA PTSD Programs Using Clinical Assessment Data

A special subcomponent of the National Mental Health Program Performance Monitoring system was designed to monitor outcomes of inpatient programs that provide specialized treatment for veterans with war-related Posttraumatic Stress Disorder.[15] Those programs were designated for ongoing, intensive performance evaluation because of the high priority placed on treatment of war-related PTSD in VA and because these programs are characterized by high treatment intensity and cost.[16] Sixty-two programs participated in this outcomes monitoring effort, which was initiated in March, 1993.

Patients who were admitted to those programs are assessed with a brief, standardized, self-report questionnaire at the time of admission, and, again, 4 months after discharge. Aperiod of four months was chosen as the follow-up interval because previous studies had shown that almost all change in symptoms and functioning occurs during the first 4 months of outpatient treatment following discharge.[16] Questionnaires ware administered either in face-to-face encounters or, when necessary, over the telephone; however, all data are based on patient's self reports (ie, the interviewers make no judgement-based ratings).

Sample

Between March 1, 1993, and February 29, 1996, 8,195 veterans were enrolled in the monitoring protocol. Admission data were successfully merged with administrative data for 6,639 veterans (81% of the total); and 4,859 (59%) were successfully contacted after discharge and completed the follow-up interview. Complete performance assessment data including both administrative indicators and the follow-up questionnaire were available for 4,165 veterans (51%). Table 1 presents comparisons of baseline assessment data from entrants with complete and incomplete data. Multiple logistic regression comparing veterans in these groups showed that, after adjusting for other measures listed in Table 1, veterans with complete data fit one of the following: less likely to be black ($P < 0.0001$); more likely to be married ($P < 0.0001$) and working ($P < 0.0012$); and lived more near to the program in which they were treated ($P < 0.03$). There were no differences on baseline clinical status measures.

Among those with successful matches in the administrative data files (n = 6,631), Table 2 compares veterans who were not reinterviewed (n = 2,474, 37%) with those who were (n = 4,165, 63%). Those who were interviewed again were more intense users of VA services and had significantly more hospital days during the 6 months before discharge and more outpatient contact after discharge.

Measures

Administrative Performance Measures. To the extent possible, administrative quality measures were based on quality measures used in the Health Plan Employer Data and Information System (HEDIS).[8] The measures of inpatient care include the following: (1) average bed days of care in psychiatric programs during the 6 months *before* discharge from the index stay; (2) average bed days of care in the 6 months *after* discharge; (3) number of readmission episodes during the first 6 months after discharge; (4–6) readmission rates at 14, 30, and 180 days; and (7) the number of days from discharge to the first readmission among those who were readmitted. Measures of the quality of follow-up outpatient care include the following: (8) the proportion of discharged inpatients who had any psychiatric outpatient visit during the first 30 days after discharge; (9) the number of outpatient visits during the first 6 months after discharge among those who had a visit during that time; (10) the average number of days from discharge to the first outpatient visit; (11) a measure of continuity of care, which was the average number of 2-month periods during the first 6 months following discharge in which a patient had two or more outpatient mental health visits; and (7) the proportion of dually diagnosed patients who had at least one psychiatric and one substance abuse clinic visit during the 6 months after discharge.

TABLE 1. Veteran Characteristics for Those With Incomplete Data Versus Those With Complete Data (ie, a completed follow-up interview and a match with administrative data bases)

	Data Incomplete (n = 4,030)		Complete Data (n = 4,165)					
	Mean/n	SD/%	Mean/n	SD/%	F	χ^2	df	P
Sociodemographic characteristics								
Age	46.75	4.80	47.40	5.17	34.84		1, 8235	0.0001
Black	1,116	27.7%	829	19.9%		68.69	1	0.0001
Hispanic	203	5.0%	190	4.6%		1.01	1	0.31
Married	1,310	32.5%	1,833	44.0%		114.60	1	0.0001
Education	12.59	1.93	12.73	2.00	10.12		1, 8214	0.0015
Miles from program	256	403	199	366	9.58		1, 8280	0.002
Past incarceration (1)	1.00	0.85	0.90	0.83	28.48		1, 8195	0.0001
Currently working	814	20.2%	1,195	28.7%		79.84	1	0.0001
Service connected for PTSD	2,018	50.1%	2,087	50.1%		0.00	1	0.98
Employment earnings (monthly)	190	679	218	594	3.89		1, 8133	0.05
Health status								
Clinical diagnosis of PTSD	3,889	96.5%	4,048	97.2%		3.19	1, 7910	0.07
PTSD (short Mississippi)	41.11	5.51	40.95	5.47	1.80		1, 8193	0.18
NEPEC PTSD scale	17.18	2.32	17.12	2.28	1.18		1, 8187	0.28
Alcohol index (ASI)	0.22	0.23	0.20	0.22	26.21		1, 8184	0.0001
Drug index (ASI)	0.10	0.11	0.08	0.10	25.82		1, 8181	0.0001
Violence	1.91	1.39	1.83	1.38	7.65		1, 8188	0.0057

Clinical Outcome Measures. Clinical outcomes were assessed in the following four domains: (1) PTSD symptoms; (2) substance abuse; (3) violent behavior; and (4) employment earnings. PTSD symptoms are measured in two ways, either by using the Short Form of the Mississippi Scale for Combat-Related PTSD, an instrument that has been validated in a large sample of outpatients[17] or by using a 4-item PTSD Scale developed at the Northeast Program Evaluation Center (the NEPEC PTSD scale) (Cronbach alpha=0.67). The NEPEC PTSD Scale correlates 0.61 and 0.74 with the Short Mississippi Scale at admission and at 4 months follow up, respectively. Those scales correlate sufficiently highly with each other and, therefore, indicate that they are measuring the same domain, but not so highly as to be redundant with each other. In an intensive outpatient PTSD study,[18] the NEPEC PTSD Scale and the Short Mississippi Scale correlated 0.63 and 0.64, respectively, with a continuous PTSD score derived from the SCID PTSD module (Structured Clinical Interview for DSM-III).[19] Additionally, in an outcome study of intensive inpatient treatment of PTSD,[16] the NEPEC PTSD Scale and the Short Mississippi Scale correlated .40 and .39, respectively, with the CAPS (Clinician Administered PTSD Scale).[20-21] Alcohol and drug abuse are measured using the composite indices from the Addiction Severity Index,[22] which is a widely used and well validated measure of substance abuse outcomes. Violent behavior is measured by 4 items that were adapted from the National Vietnam Veterans Readjustment Study, as follows[23]: (1) the destruction of property; (2) threatening someone with physical violence without a weapon; (3) threatening someone with a weapon; and (4) physically fighting with someone (Cronbach alpha = 0.71). Employment was measured by the number of dollars earned from employment during the 30 days preceding each report.

Analyses

The nalyses proceeded in several stages. In the first stage, data from the outcomes monitoring protocol were merged with administrative data and paired *t* tests were used to evaluate the significance of changes in measures from admission to follow up.

In the second stage, measures of clinical improvement were constructed using the residuals from the regression of post-discharge outcomes

TABLE 2. Administrative Indicator Data by Follow-Up Status

Administrative indicators	Merged/No Follow-Up (n = 2,474) Mean/N	SD/%	Merged/With Follow-Up (n = 4,165) Mean/N	SD/%	F	χ^2	df	P
Inpatient days: six months before disch.	58.0	39.1	54.6	37.3	12.3		1, 6688	0.0005
Inpatient days: six months after disch.	7.81	19.49	8.00	18.907	0.15		1, 6688	0.69
Readmission episodes: six months after disch.	0.42	0.76	0.41	0.72	0.49		1, 6688	0.48
Readmission within 14 days	66	2.7%	82	2.0%		3.66	1	0.056
Readmission within 30 days	121	4.9%	176	4.2%		1.78	1	0.182
Readmission within 180 days	641	25.9%	1129	27.1%		0.782	1	0.377
Days to readmission among readmitted	84.2	50.3	89.8	50.2	5.1		1, 1768	0.0243
Any outpatient visit within 30 days of discharge	1389	56.1%	2716	65.2%		47.24	1	0.0001
Outpatient visits within 6 months of discharge	14.5	20.5	18.5	24.2	41.73		1, 5983	0.0001
Days to first outpatient visit within 6 months of DC	32.6	37.1	29.0	35.5	14.1		1, 5973	0.0002
Continuity of care	1.59	1.18	1.90	1.17	109.2		1, 6670	0.0001
Psych and SA treatment for dually diagnosed	330	13.3%	587	14.1%		0.5	1	0.473

on baseline measures. Covariates included age, race, marital status, and the baseline values of the 6 clinical outcome measures.

The associations between the two types of measures were then evaluated using hierarchical linear modeling (HLM).[24] HLM is a technique that may be applied to data in which individual measurements cannot be considered independent because they are clustered within larger groups, eg, in patients clustered within distinct VA hospitals. One must use this type of modeling rather than standard linear regression because the correlation among individual patient observations will bias variance estimates and augment Type I error. The PROC MIXED procedure of the SAS (R) software system (SAS Institute, Cary, NC) was used for all analyses.

Two sets of analyses were conducted. The first set measured the association of administrative quality measures, such as the independent variables, and outcomes, such as the dependent variables, at the level of individual patients. Random intercept and random slope components were included in the model to account for similarity of patient outcomes, within a given hospital, that is attributable to underlying differences between hospitals. This was necessary because, given the same administrative quality measures, outcomes may be more similar among patients treated at the same hospital than between patients treated at different hospitals. The same set of patient covariates used to create the adjusted outcome measures were also used to adjust the administrative measures.

The second set of analyses measured the association between administrative quality measures and outcomes at the level of individual VA medical centers. In those hierarchical models, which are sometimes called "means-as-outcomes" models,[24] the random intercept represents the hospital-level mean of its patient outcomes and is modeled at the second level of the hierarchy as a function of hospital quality means plus random error.

Results

Sample Characteristics

Sociodemographic and clinical characteristics of the sample and of their representations are pre-

TABLE 3. Clinical Improvement Scores

	Baseline	SD	4 Months	SD	diff	t*	P
PTSD (short Mississippi)	41.2	5.5	38.9	7.0	−2.4	17.06	<0.0001
NEPEC PTSD scale	17.2	2.3	16.1	3.0	−1.1	19.16	<0.0001
Alcohol index (ASI)	0.22	0.23	0.16	0.19	−0.1	8.87	<0.0001
Drug index (ASI)	0.09	0.11	0.07	0.08	−0.0	8.27	<0.0001
Violence	2.0	1.4	1.2	1.3	−0.7	23.67	<0.0001
Employment income	218.54	594.40	175.66	552.26	−42.88	4.59	<0.0001

*Paired *t* tests.

sented in Tables 1 and 2. Table 3 shows that significant improvement was observed on 5 of 6 outcome measures ($P < 0.0001$). The one measure of employment income showed a significant decline ($P < 0.0001$).

Table 4 presents HLM analyses of the association of administrative indicators and improvement at the level of individual patients. Five of twelve administrative measures were significantly associated with clinical outcome measures. Most importantly, the number of readmission episodes that occurred during the 6 months after discharge was associated with poorer outcomes on 5 of 6 outcome measures. Both the number of days hospitalized and the occurrence of at least one readmission during the first 180 days after discharge were correlated with poor outcomes on 4 improvement measures, both PTSD symptom scales, violence, and employment. The fact that the outpatient visit was within 30 days of discharge and that the number of outpatient visits during the 6 months after discharge were associated with reduced levels of alcohol problems. The average number of days to re-hospitalization among those rehospitalized was associated with higher violence scores, which suggests that among people who were rehospitalized those who delayed admission may have had more violent behavior.

There were no significant associations between administrative measures and outcomes at the level of VA medical centers.

Discussion

Principal Findings

This study found a substantial number of statistically significant correlations between administrative performance indicators and clinical outcomes. Indicators that were related to hospital readmission data were the most strong indicators of clinical outcome as compared with measures of access, intensity, or continuity of outpatient care.

Although readmission is one of the most commonly used performance indicators in behavioral health, this study represents only the second empirical examination of the relationship of this or any other indicator to clinical outcomes as measured by psychometrically standardized instruments. In the other study of this type, Lyons et al[25] did not find any relationship between clinical status at the time of discharge from a hospital episode and readmission within 30 or 180 days. The most likely explanation for the difference in results between that study and ours is the timing of the clinical outcome assessment. Lyons et al assessed outcomes at the time of hospital discharge and evaluated their ability to predict readmission. In contrast, in this study outcomes were assessed 4 months after discharge during the period of risk for readmission. We feel that outcomes that were used to validate performance measures were most informative if they reflect post-discharge health status; that is because the goal of inpatient treatment is to improve clinical status after return to life in the community, not while the patient is still under hospital care.

It is also notable that, in this study, the total number of readmission episodes was a more strong and consistent indicator of adverse clinical outcome than was the simple readmission at various fixed time intervals, which was the more commonly used type of measure. Mental health program managers may want to consider using this new measure instead of, or in addition to, the more traditional indicators.

TABLE 4. Relationship of Administrative Indicators and Four-Month Outcomes*

	Clinical Improvement Measures					
Administrative Indicators	PTSD: Short Mississippi	PTSD: NEPEC Scale	Alcohol Problems (ASI)	Drug Problems (ASI)	Violent Behavior	Employment Income
Inpatient days: six months before discharge	−0.0098 0.5747	−0.0415 0.1451	−0.0078 0.6886	−0.0082 0.6386	−0.0098 0.6044	−0.0253 0.3809
Inpatient days: six months after discharge	**0.0526** **0.0023**	**0.0717** **0.0001**	0.0228 0.187	0.0183 0.2905	**0.0689** **0.0001**	**−0.0521** **0.0028**
Readmission episodes: six months after discharge	**0.072** **0.0001**	**0.0807** **0.0001**	**0.0563** **0.0014**	0.0249 0.1569	**0.0688** **0.0001**	**−0.0463** **0.0093**
Readmission within 14 days	−0.0171 0.3245	0.0217 0.2436	−0.0189 0.2761	0.0154 0.3827	−0.008 0.7371	0.0089 0.6063
Readmission within 30 days	−0.0215 0.2162	0.0365 0.0523	0.0274 0.19	0.0216 0.3312	−0.0119 0.5508	−0.0243 0.1839
Readmission within 180 days	**0.064** **0.0027**	**0.0685** **0.0004**	0.0354 0.0628	0.034 0.1108	**0.0753** **0.0001**	**−0.0533** **0.0039**
Days to readmission among those readmitted	0.0215 0.5186	−0.0484 0.1556	−0.0108 0.7684	−0.0264 0.5121	**0.0984** **0.0173**	0.0081 0.7943
Any outpatient visit within 30 days of discharge	−0.02441 0.2411	0.0059 0.7545	**−0.0407** **0.0369**	−0.009 0.6025	−0.0362 0.0908	−0.0117 0.5167
Outpatient visits within 6 months of discharge	0.0084 0.6776	0.0253 0.1642	**−0.0549** **0.0085**	−0.0192 0.3727	0.0021 0.9253	0.0134 0.463
Days to first outpatient visit 6 months after discharge	0.0371 0.086	0.0037 0.8573	0.0334 0.1331	0.0156 0.3758	0.0238 0.2282	−0.0003 0.9875
Continuity of Care (bi-months with 2 or more OP visits)	0.0069 0.761	0.0304 0.0836	−0.0293 0.0953	0.0067 0.7156	−0.0002 0.9929	−0.0324 0.0663
Psych and SA treatment for dually diagnosed	−0.0096 0.5819	0.013 0.5065	0.0153 0.5126	0.0019 0.9255	−0.0153 0.4302	−0.0006 0.9861

*Standardized regression coefficients (above) and P value (below).
Bolded values $P < 0.05$.

The lack of significant relationships at the level of hospitals may reflect the fact that there is less variability between hospitals than between individual patients; therefore, significant relationships are less likely to be detected.

Implications for Health Care Quality Evaluation

There is now widespread interest in evaluating clinical outcomes in large health care systems.[7] Those outcomes are regarded by many as the "gold standard" measure of health care value. From that perspective the low correlations between administrative indicators and outcome measures might seem to argue against continued use of those indicators. We believe, however, that although the ultimate goal of healthcare is to optimize healthcare, the outcomes maximization should not be regarded as the sole or ultimate indicator of health care quality in performance monitoring systems. As Donabedian[26] emphasized decades ago, health care quality assessment must consider the three dimensions of care of structure, process, and outcome.

In this broader view, even when outcome measures show good treatment results, we would be concerned about quality of care if large numbers of patients discharged from the hospital did not receive follow-up care or if patients who did receive follow-up care had to wait 5 months before their first visit. Even when they are responding well to treatment, patients need to be followed closely to maximize the likelihood that those gains are maintained. The "tyranny of outcomes" can lead to an overemphasis on one dimension of quality at the expense of attention to quality in other areas.[27,28]

Although their relationship to clinical outcomes is weak, administrative indicators also have to face the validity of their own in addressing issues of clinical process. They are often of greater practical use than outcome measures because they more clearly suggest specific remedial actions. For example, whereas a facility with poor outcomes on PTSD symptom scales may have little idea of how to improve those outcomes, a provider whose patients do not get to their outpatient treatment in as timely a fashion as those from other facilities clearly needs to re-evaluate referral procedures.

It is also important to note that the cost of assessing outcomes by interviewing patients at baseline and after a specified follow-up period is substantial. We estimate that the per patient cost of the outcome data presented here is approximately $150 to $300 per patient. Applying this estimate to the entire VA mental health program, which treats approximately 580,000 patients per year, comprehensive outcomes monitoring would result in a total cost of $87 to $174 million per year, which is a 4% to 8% increase in the entire VA mental health budget. Data for administrative indicators, in contrast, are far less expensive, and costs an estimated $0.25 per patient, which includes both labor and data processing costs, 0.06% of current VA mental health expenditures. Thus, in addition to being significantly related to clinical outcomes, administrative indicators have face validity as clinical process measures and offer an efficient way of monitoring performance on a large scale. Whereaa this study shows that they can not be used as substitutes for clinical outcome measures, they retain an important complementary role. More costly outcomes measures are best gathered on smaller cohorts of patients selected on the basis of their special salience for particular institutions.

Limitations

Several limitations of this study must be commenetd on. First, although a strength of this study from the scientific point of view is that the sample is diagnostically homogeneous, this limits the generalizability of the findings to other diagnostic groups.[29] Our findings may also not apply to health care systems other than VA.

Second, administrative data on outpatient care are only available from VA clinics and do not capture care provided through referral to non-VA clinics. That limitation is intrinsic to administrative data sets and may partially explain the low correlations between outpatient performance and clinical outcomes. Previous studies, however, have shown that non-VA treatment is quite limited in this population.[30]

Third, our examination of the relationship of administrative indicators and clinical outcome measures involved a total of 72 analyses raising the possibility that significant findings were a spurious artifact of the large number of comparisons. With 72 comparisons, on the chance that we would expect only 4 significant correlations at $P < 0.05$. This, however, is only one-fourth of the 16 significant relationships we observed. The Bonferroni correction, a far more conservative method of adjusting for multiple comparisons, would reduce the alpha level required for significance to $P < 0.0006$. Even such a low alpha would not substantially change our principal results.

Fourth, complete data for our analyses were available for only 51% of the original sample and the clear selection biases were identified in our analysis of baseline data from all entrants. As the goal of this study was to examine the relationship of different types of outcome measures to one another, this selection bias is unlikely to affect the validity of our findings. Our follow-up rates were similar to those obtained in other projects of this type.[31]

Finally, it should be noted that some of the clinical outcome measures used were specifically shortened and adapted to the demands of large scale outcome monitoring effort. Test-retest reliability data were not available from this sample, and we have no information on variability in the administration of the questionnaire. Methods for addressing this variability through training and ongoing reliability testing would not be difficult to design, but would be costly to implement.

Administrative performance measures are significantly associated with outcomes at the level of individual clients but not at the level of facilities. Whereas the significant correlations observed here tend to validate the administrative data, both types of data are needed in to evaluate of health system performance.

References

1. **Edmunds E, Frank R, Hogan M, McCarty, Robinson-Beale, Weisner C, eds.** Managing managed care: Quality improvement in behavioral health. Washington, DC: National Academy Press, 1997.

2. **Goleman D.** Critics say managed care eroding mental care. New York Times, January 24, 1996, C9.

3. **Callahan JJ, Shepard DS, Beinecke RH, Larson MJ, Cavanaugh D.** Mental health/substance abuse treatment in managed care: The Massachusetts medicaid experience. Health Affairs 1995;14(3):173.

4. **Christianson J, Manning W, Lurie N, Stoner TJ, Gray DZ, Popkin M, et al.** Utah's prepaid mental health plan: The first year. Health Affairs 1995;14(3):160.

5. **Mechanic D, Schlesinger M, McAlpine DD.** Management of mental health and substance abuse services: State of the art and early results: The Milbank Quarterly 1995;73(1):19.

6. **Wells K, Astrachan BM, Tischler GL, Untzer J.** Issues and approaches in evaluating managed mental health care. The Milbank Quarterly 1995;73(1):57.

7. **Sederer L, Dickey B.** Outcomes Assessment in Clinical Practice. Baltimore: Williams and Wilkins, 1996.

8. **National Committee on Quality Assurance (NCQA).** Report card pilot project. Washington, DC: The National Committee on Quality Assurance, 1995.

9. **Cleary PD, Edgman-Levitan S, Roberts M, Moloney TW, McMullen W, Walker JD, et al.** Patients evaluate their hospital care: A national survey. Health Affairs 1991;10:254.

10. **Allen HE, Darling H, McNeil DN, Bastien F.** The employee health care value survey: Round one. Health Affairs 1994;13(1):25.

11. **Rosenheck RA, Cicchetti D.** A mental health program report card: A multidimensional approach to performance monitoring in public sector programs. Community Mental Health Journal 1998;34(1):85.

12. **Rosenheck R.** Department of veterans affairs national mental health program performance monitoring system: Fiscal 1997 report. West Haven: Northeast Program Evaluation Center, 1998.

13. **Rosenheck RA, Wilson N, Meterko M.** Consumer satisfaction with inpatient mental health treatment: Influence of patient and hospital factors. Psychiatric Services 1997;48:1553.

14. **Rosenheck R, Fontana A, Spencer H, Grey S.** Long journey home V: Treatment of posttraumatic stress disorder in the Department of Veterans Affairs: Fiscal year 1996 service delivery and performance. West Haven:Northeast Program Evaluation Center, 1997.

15. **Fontana A, Rosenheck R.** Outcome Monitoring of Specialized Intensive PTSD Programs: Fiscal Year 1996 Report. West Haven: Northeast Program Evaluation Center, 1997.

16. **Fontana A, Rosenheck RA.** Effectiveness and cost of inpatient treatment of posttraumatic stress disorder, American Journal of Psychiatry. 1997;154:758.

17. **Fontana A, Rosenheck R.** A short form of the Mississippi Scale for measuring change in combat-related PTSD. J Traumatic Stress 1994;7:407.

18. **Fontana A, Rosenheck R, Spencer H.** The long journey home III: The third progress report on the specialized PTSD programs. Wewst Haven: Northeast Program Evaluation Center, 1993.

19. **Spitzer RL, Williams JBW.** Structured clinical interview for DSM-III PTSD. New York: New York State Psychiatric Institute, 1985.

20. **Blake DD.** Rationale and development of the clinician-administered PTSD scales. PTSD Research Quarterly 1994;5(2):1.

21. **Weathers FW, Litz BT.** Psychometric properties of the clinician-administered PTSD scale, CAPS-1. PTSD Research Quarterly 1994;5(2):2.

22. **McLellan AT, Luborsky L, Cacciola J, Griffith J, Evans F, Barr HL, et al.** New data from the Addiction Severity Index: reliability and validity in three centers. J Nervous Mental Dis 1985;173:412.

23. **Kulka RA, Schlenger WE, Fairbank JA, Hough RL, Jordan BK, Marmar C, et al.** Trauma and the Vietnam war generation: Report of findings from the national Vietnam veterans readjustment study. New York: Brunner/Mazel, 1990.

24. **Bryk AA, Raudenbush SW.** Hierarchical linear models. New bury Park: Sage Publications, 1992.

25. **Lyons JS, O'Mahoney MT, Miller SI, Neme J, Kabat J, Miller F.** Predicting readmission to the psychiatric hospital in a managed care environment: Implications for quality indicators. Am J Psych 1997;154:337.

26. **Donabedian A.** Evaluating the quality of medical care. Milbank Memorial Fund Quarterly 1996;44:166.

27. **Berwick DM.** Toward an applied technology for quality measurement in health care. Med Decis Making 1988;8:253.

28. **Hammermeister KE, Shroyer L, Sethi GK, Grover FL.** Why is it important to demonstrate linkages between outcomes of care and processes and structures of care. Med Care 1995;33(Suppl):OS5.

29. **Druss B, Rosenheck RA.** Evaluating behavioral health care quality: An empirical study of HEDIS. Psych Services 1997;48:71.

30. **Fontana AF, Rosenheck RA.** Improving the efficiency of resource utilization in outpatient treatment of post-traumatic stress disorder: A report on specialized PTSD programs. West Haven: Northeast Program Evaluation Center, 1994.

31. **Steinwachs D, Wu A, Skinner E.** How will outcomes management work? Health Affairs 1994; 13:153.