Acceptability and effectiveness of patient–reported assessments and feedback in a managed behavioral health care setting

Article in The American Journal of Managed Care - January 2006

CITATIONS 33

READS 223

7 authors, including:

Benjamin B Brodey
TeleSage, Inc., Chapel Hill, NC, United States
23 PUBLICATIONS 266 CITATIONS

Mark Maruish
27 PUBLICATIONS 758 CITATIONS

Inger Brodey
University of North Carolina at Chapel Hill
19 PUBLICATIONS 149 CITATIONS

Jürgen Unützer
University of Washington Seattle
264 PUBLICATIONS 19,257 CITATIONS

Some of the authors of this publication are also working on these related projects:

Validation of the NetSCID View project

T-ASI Teen Addiction Severity Index for Web and IVR View project
Since Howard and colleagues introduced the paradigm of patient-focused research, investigations of patient-reported assessments (like those solicited by the 2003 National Institutes of Health Request for Applications, “Dynamic Assessment of Patient-Reported Chronic Disease Outcomes”) have focused primarily on their usefulness as a research tool, tracking large groups to assess the effectiveness of a particular treatment or to streamline processes for cost control. However, a growing body of research suggests that a reliable patient-reported assessment tracking system may offer direct clinical benefits as well.\(^2\) This may be particularly true in psychotherapy, where studies suggest that a lack of reliable feedback mechanisms may contribute to therapists’ inability to make accurate assessments of patient progress.\(^5\)\(^,\)\(^6\) Similarly, in a meta-analysis of available studies, Sapyta et al\(^3\) found that when patients provide feedback to psychologists regarding the patients’ general health and progress in treatment, significant benefits may result, especially for patients who are not doing well. Brown and Jones\(^7\) suggest that early identification of patients who are not progressing in treatment can help clinicians keep patients engaged in the process long enough to improve.

Lambert and colleagues\(^8\)\^-\(^10\) have conducted a series of studies to investigate whether providing periodic feedback to therapists about patient progress leads to improvement in treatment outcomes. Their results showed that feedback to therapists resulted in fewer treatment sessions with no worsening of outcome for patients making progress, while patients who were deteriorating or not progressing received more sessions and had better outcomes than predicted. Patient-reported feedback, then, allows therapists to focus resources where needed, ensuring that those patients most at risk of failing to respond to treatment receive the attention and encouragement they need to continue in treatment.

Even a minimal feedback program can have real benefits for patients and therapists, increasing the effectiveness of therapy. In a 2002 study, Percevic showed that 1-time feedback provided to therapists early in treatment can yield significant effects for all patients, presumably by allowing therapists to more closely tailor treatment programs to patient needs.\(^1\)\(^,\)\(^1\) Early feedback also can help therapists identify those patients with more severe symptoms and initiate early efforts to keep those patients in treatment.

An effective patient-reporting tool, then, can have real clinical applications, provided it can be administered economically and feedback can be provided in a timely manner. Such a tool could help therapists to tailor the treatment to the patient, based on real-time information about a patient’s current status and improvement over time. In short, automated clinical feedback from patient surveys can allow clinicians to “focus on what they do...”
best: forming relationships with patients, weighing the available information, making mutually informed treatment decisions, and conducting the appropriate therapy." Ideally, such a system would result in improved outcomes for the individual patient while reducing time in treatment and, consequently, the cost of treatment.

As the providers of mental health services for approximately 75% of the 180 million insured Americans, managed behavioral health organizations (MBHOs) have a vested interest in tracking care and allocating resources effectively. Managed behavioral health organizations have long been in the business of tracking outcomes for large groups and are now beginning to focus on individual outcomes as well. TeleSage, an automated survey and health outcomes tracking company, and United Behavioral Health (UBH), a national MBHO, collaborated to develop a randomized, longitudinal, outcomes-tracking study of UBH's enrollees. The object of the study was to determine whether providing clinicians with the results of mental health assessments at intake would have a significant impact on patients' mental health outcomes and provider satisfaction at 6 weeks.

METHODOLOGY

Patients recruited into the study were asked to complete 11 items from the Symptom Checklist-90 (SCL-90) both when they were authorized to receive services through UBH and at 6 weeks after authorization. No feedback reports were generated for half of the subjects; for the remainder, reports detailing survey results were generated after the initial and 6-week administrations and sent to the clinician providing treatment.

Shortly after the second feedback report was sent at 6 weeks after intake, satisfaction surveys were mailed to clinicians who had been sent the reports. The surveys asked clinicians whether they remembered receiving and reading the reports, and asked them to offer opinions regarding the usefulness, clinical relevance, and user-friendliness of the report formats. The surveys were designed to assess the usefulness of the reports and to gather suggestions for improving the reports for clinical application.

The SCL-11

The SCL-11 was developed for this study; a list of the SCL-90 items used is available from the primary author of this article (BBB). Intended as a measure of depression and anxiety, the SCL-11 includes 11 items taken verbatim from the depression (5 items) and anxiety (6 items) subscales of the SCL-90. Items were selected based on their predictive value relative to the full depression and anxiety domains of the SCL-90 and for their sensitivity to change. Independent validation of the SCL-11 was conducted prior to beginning this study. Internal consistency (depression, Cronbach's $\alpha = .90$; anxiety, Cronbach's $\alpha = .78$ for a nonpatient sample) and test-retest reliability (depression, Pearson $r = .61$; anxiety, Pearson $r = .61$ for a patient sample) for the depression and anxiety measures are comparable to those of the SCL-90.

Clinician Reports

Clinicians received an initial summary report that offered graphical, numeric, and written interpretations of the patient's responses to the SCL-11 items and an explanation of the interpretation of scores. In addition, the report included the text of questions for which the patient gave extreme responses and information on the time needed to complete the survey. This report included a simple bar graph comparing the patient's numeric scores with standardized population norms. Clinicians also received a report at 6 weeks, which replaced the bar graph with a longitudinal graph representing patient progress over time, intended to allow clinicians to gauge deterioration or improvement. The format of the final report required 3 pages. In addition, 4 introductory pages explained the project to clinicians, resulting in an overall length of 7 pages. To protect confidentiality, reports were identified only by a UBH-assigned patient identification number.

Recruitment and Attrition

The study population included UBH patients who were authorized to receive outpatient mental health services from a network provider. Subjects represented a cross-section of UBH enrollees and dependents seeking mental health treatment. United Behavioral Health is contracted to administer mental health benefits nationwide; its members are predominantly urban, though not inner-city, residents of all 50 states. The majority is private pay and is not Medicare or Medicaid recipients. Participants were 18 years or older at the time of their request for authorization to receive services, not at risk of harming themselves or others or otherwise in need of an emergency intervention, and not cognitively impaired.

Two recruitment methods were used. First, 10 UBH call-center intake coordinators recruited patients to participate in the project by introducing the study to consecutive, eligible plan members after the individual's needs had been addressed and an authorization for care had been given. If the patient expressed an interest in participating, the intake coordinator transferred him or her to an automated interactive voice response (IVR)
The IVR system presented information about the study, including a brief description of its purpose and procedure, as well as information about potential risks, confidentiality, and patient rights. Patients also were provided with the names and phone numbers of the study coordinator and a UBH patient advocate. Patients could provide consent to participate by pressing a touch-tone button in response to a request for informed consent. This procedure was approved by Western Institutional Review Board. Once informed consent had been given, the IVR system immediately administered the SCL-11.

To ensure an adequate study population, we also recruited participants via US mail. In a pilot study, we demonstrated that surveys administered by IVR and paper and pencil yield similar results among both English- and Spanish-speaking subjects. The SCL-11 and a copy of the informed-consent script were mailed to all eligible patients who had authorized UBH, as a matter of routine, to contact them by mail. (A small number of patients request that UBH not mail materials to them, usually because of fears regarding confidentiality.) To protect confidentiality, all mail was sent in envelopes that had only a post office box for the return address and did not identify UBH as the sender. Those interested in participating were asked either to complete the informed-consent document and SCL-11 and return them via US mail, or to dial into the IVR system.

To increase follow-up response rates and minimize attrition, the SCL-11 was administered at 6 weeks using procedures adapted from Dillman, including presurvey notification letters, thank-you letters, reminder letters, and telephone follow-up calls to nonresponders. At the follow-up assessment, participants were given the option of completing the SCL-11 either via IVR using a toll-free telephone line or by returning a paper-and-pencil version of the survey via US mail.

The project was approved by an institutional review board. After a brief description of the study was provided, informed consent was obtained from all subjects, either in the form of a signed document or via the IVR mechanism described above. No incentives were offered for participation.

Analysis

Statistical analyses compared changes in group scores derived from the outcome scales. Subject outcomes were analyzed based on intention to treat; that is, data for participants who completed the follow-up assessment were included regardless of whether they actually received treatment. Change scores for those participants who completed follow-up surveys were derived by subtracting the baseline score for each domain from the 6-week outcome measures. The GLM procedure (SAS/STAT software, SAS Institute Inc, Cary, NC) was used to construct analysis-of-covariance (ANCOVA) tests of the effect of clinician feedback.

Clinician responses to the satisfaction survey were analyzed for the relationship between patient improvement and clinician satisfaction. Secondary statistical analyses examined responses of subgroups of clinicians based on age, education, and experience with outcomes measures.

RESULTS

Sample Characteristics and Attrition Bias

A total of 1374 patients aged 18 years or older were enrolled in the study. The demographics of subjects who agreed to participate were compared with those of UBH enrollees seeking authorization for treatment at the time the study was conducted. Subjects were slightly more likely to be female and to obtain mental health services, although no statistically significant differences were noted. Approximately 20% of subjects did not receive treatment during the 6 weeks between their initial and follow-up assessments.

The great majority of the study population—87.5%—was white; 4.5% identified themselves as black and another 4% said they were Hispanic. The remaining 4% said they were multiracial or belonged to other ethnic groups. Twenty-seven percent of participants were male. Participants were randomly assigned to 1 of 2 conditions: clinician feedback or no feedback. Although some differences were noted for age (P = .03, \( \chi^2(3) = 8.62 \); see Table 1) and patient's relationship to the insured member (P = .02, \( \chi^2(2) = 8.27 \); see Table 1), chi-square tests comparing subject characteristics across the groups did not otherwise approach significance, indicating that randomization produced equivalent groups.

Similarly, analysis of variance comparing the baseline status of each outcome measure across the feedback groups indicated equivalent levels of distress and dysfunction of study participants at entry into the study. Depression was noted as a diagnosis on at least 1 claim for 40% of subjects; anxiety was noted for 15%. Attrition was within the expected range for a study for which assessment measures were integrated with routine clinical practice. Of the 1374 participants completing the baseline assessment, 954 (69%) completed the 6-week assessment. The remaining 31% could not be reached for follow-up. The rate of attrition across groups was comparable between baseline and the 6-week follow-up (\( \chi^2 = 6.1, P = .11 \)).
Because of the possibility of attrition bias, we examined whether participant characteristics, including responses to SCL-11 items as well as demographic indicators, predicted attrition using bivariate tests of association (t tests and chi-square tests). No characteristic predicted completion of the follow-up measure at 6 weeks. This analysis suggests that comparisons of the study cohorts on the main outcome measures were unlikely to be biased by differential dropout of members with more severe distress or dysfunction. The only potential threat involved dropout of members who did not receive any mental healthcare after their entry into the study in the feedback group. Because respondents in this group generally had lower rates of treatment, any observed outcome advantages attributable to clinician feedback were achieved despite these lower overall rates of treatment.

Clinical Benefits

Group change scores derived from the outcome scales and ANCOVAs are summarized in Table 2. Controlling for age and the patient’s relationship to the insured member (the only factors with P values significant at the .10 level and with a confounding effect), the results at 6 weeks showed a small but statistically significant effect of clinician feedback on patient outcome measures. Patients of clinicians who received the baseline feedback report showed greater improvement in mean domain scores for total symptoms (0.35 vs 0.25 on a 5-point Likert scale) than patients of clinicians who did not receive feedback. This difference constitutes a 28% greater improvement with clinician feedback compared with no feedback. Similarly, mean domain improvement was greater in the clinician feedback group for depression (0.41 vs 0.29; 29% greater improvement).

Clinical Utility and Satisfaction

Clinician satisfaction surveys were sent to the 691 providers who received reports. Clinicians who received reports on more than 1 patient in the study were sur-

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No Feedback (n = 693)</th>
<th>Feedback (n = 681)</th>
<th>( \chi^2 )</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age 18-29</td>
<td>180</td>
<td>136</td>
<td>8.62</td>
<td>&lt;.03</td>
</tr>
<tr>
<td>30-39</td>
<td>222</td>
<td>225</td>
<td>3.3</td>
<td>.08</td>
</tr>
<tr>
<td>40-49</td>
<td>180</td>
<td>184</td>
<td>2.7</td>
<td>.11</td>
</tr>
<tr>
<td>50+</td>
<td>111</td>
<td>136</td>
<td>2.0</td>
<td>.16</td>
</tr>
<tr>
<td>Insurance relationship</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insured</td>
<td>503</td>
<td>449</td>
<td>8.27</td>
<td>&lt;.02</td>
</tr>
<tr>
<td>Spouse</td>
<td>163</td>
<td>207</td>
<td>2.5</td>
<td>.12</td>
</tr>
<tr>
<td>Dependent</td>
<td>27</td>
<td>25</td>
<td>0.4</td>
<td>.53</td>
</tr>
</tbody>
</table>

**Table 1.** Age and Insurance Relationship of Participants

<table>
<thead>
<tr>
<th>Symptom</th>
<th>Score at Intake, Mean ± SD</th>
<th>Score at 6 Weeks, Mean ± SD</th>
<th>ANCOVA Results for Feedback</th>
<th>F Test</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No Feedback (n = 693)</td>
<td>Feedback (n = 681)</td>
<td>No Feedback (n = 487)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td>1.01 ± 0.925</td>
<td>0.96 ± 0.933</td>
<td>0.72 ± 0.797</td>
<td>0.21</td>
<td>0.27</td>
</tr>
<tr>
<td>Depression</td>
<td>1.60 ± 1.026</td>
<td>1.60 ± 1.042</td>
<td>1.26 ± 0.987</td>
<td>0.29</td>
<td>0.41</td>
</tr>
<tr>
<td>Total symptoms</td>
<td>1.33 ± 0.884</td>
<td>1.31 ± 0.897</td>
<td>1.02 ± 0.830</td>
<td>0.25</td>
<td>0.35</td>
</tr>
</tbody>
</table>

ANCOVA indicates analysis of covariance.
veyed only about their first patient. A total of 488 (70%) clinicians responded. Of these, 130 were removed from the analysis because they indicated that they had never seen the patients for whom they received reports; these patients had completed the initial assessment and been assigned to the care providers to whom their reports were sent, but then had either not sought care or had seen another provider. This left 358 valid surveys for analysis. Of this group, 74.4% recalled receiving the report; a full 97.6% of those who remembered receiving the report said that they had read it. Only 13.6%, however, indicated that they received the report prior to the patient’s initial appointment.

Survey data are summarized in Table 3. The data indicate that most clinicians found the report easy to understand, but as a group they were divided as to its utility. Written comments suggest that clinicians who reacted negatively toward the reports often were concerned with additional paperwork and feared that managed care companies were intruding on the treatment process. It is worth noting that 47% of clinicians said that the longitudinal tracking provided in the second report “helped me monitor changes in my patient over time.”

Clinician responses were compared with patient data from the SCL-11. Clinicians were more likely to say that the reports helped them provide better care when the reports indicated that their patients exhibited more symptoms of depression and when patients assessed their sense of wellness and health as being low. Similarly, clinicians were more likely to agree that the report was an aid to treatment when patients endorsed more symptoms of depression and anxiety.

No significant response differences among the various subgroups were found on any of the survey items.

**DISCUSSION**

A number of factors may account for the significant benefits of clinician feedback reports on patient outcomes at 6 weeks. Most clinicians had access to the initial SCL-11 feedback report, showing the patient’s general mental health status, well before the 6-week follow-up survey. Information in the report may have aided clinicians in assessing the patient’s condition and provided an opportunity to tailor treatment through conversation with the patient.

Many clinicians reported that they received feedback reports for patients whom they did not see. These patients may have delayed seeking care, decided not to seek care at all, or seen a different clinician. Because our analysis was based on intention to treat, patient choices like these would have diminished the observable effect of feedback in our study.

In addition to these patient factors influencing the success of clinician feedback, several problems in study implementation likely contributed to washing out the effect. First, the national scale of the study meant that clinicians were not trained to evaluate and use the results of the SCL-11. Instead, clinicians received with the feedback report a letter that described the study and explained the report; this approach proved less than ideal, both because it was not as effective as direct instruction and because it increased the length of the reports (hence, the perceived burden on clinicians). In addition, because few clinicians had more than 1 patient participating in the study, most received only 1 report; as a result, few had the opportunity to learn from experience with the report.

Second, slower-than-expected recruitment forced the research team to use US mail for delivery of some informed-consent forms and surveys, undercutting

---

**Table 3. Summary of Responses to Clinician Satisfaction Survey (n = 260)**

<table>
<thead>
<tr>
<th>Response</th>
<th>Agree</th>
<th>Not Sure</th>
<th>Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>The information on the second report helped me to monitor changes in my</td>
<td>122</td>
<td>38</td>
<td>100</td>
</tr>
<tr>
<td>patient over time</td>
<td>47%</td>
<td>14%</td>
<td>39%</td>
</tr>
<tr>
<td>The report saved me time</td>
<td>21</td>
<td>65</td>
<td>174</td>
</tr>
<tr>
<td></td>
<td>8%</td>
<td>25%</td>
<td>67%</td>
</tr>
<tr>
<td>The summary information on the first page was useful</td>
<td>152</td>
<td>49</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>58%</td>
<td>19%</td>
<td>23%</td>
</tr>
<tr>
<td>The report was easy to understand</td>
<td>217</td>
<td>15</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>83%</td>
<td>6%</td>
<td>11%</td>
</tr>
<tr>
<td>I would like to receive reports on all my patients</td>
<td>91</td>
<td>61</td>
<td>208</td>
</tr>
<tr>
<td></td>
<td>35%</td>
<td>23%</td>
<td>42%</td>
</tr>
</tbody>
</table>
the primary advantage of the IVR system that was initially designed for the study. Results collected via IVR were automatically faxed to clinicians within 24 hours of completion; those submitted on paper via US mail had to be manually entered into the system for scoring and delivery. The mailing delay, which would not have existed in the IVR model, meant that clinician reports for those patients recruited by mail frequently arrived after the patient’s first clinical appointment.

The clinician feedback reports would naturally have been most helpful to clinicians who read the reports before providing clinical treatment. Indeed, a number of researchers have noted that feedback, if it is to be effective, must be both action oriented and timely.\(^8\,19\) Unfortunately, only a small number of clinicians—13.6%—reported receiving and reading the reports before seeing the patients for the first time. Mailing delays may account for a significant portion of those clinicians who saw reports after seeing the patient.

Third, the final version of the clinician feedback report was 7 pages long. This length may have led clinicians to perceive it as complicated and burdensome to read; 66.7% disagreed with the statement, “The report saved me time,” although only 10.6% found the report difficult to understand. These perceptions may well have been influenced by the bulk of the report. More importantly, long reports may have obscured the essential information that the study was attempting to provide on patient mental health status and longitudinal change.

CONCLUSIONS

The results of this study are encouraging in that they suggest that providing clinicians with outcomes feedback has at least a short-term benefit. On the other hand, the results—and the problems encountered in study implementation—also highlight a number of potential pitfalls for individual patient outcomes tracking.

This study, which used both IVR and paper-and-pencil mental health assessments, demonstrated that the provision of clinician reports had a positive impact on subjects’ mental health status at 6 weeks. Although a number of problems—the use of US mail to collect survey responses, the lack of clinician training, the length of the reports, and the fact that some subjects never received treatment—likely lessened the impact of clinician reports, this study shows that automated mental health assessments and outcomes tracking can improve the chances for patients’ mental health recovery.

We believe that this article represents one of the first randomized studies to examine the clinical effectiveness of routine outcomes tracking using IVR. Because our analysis was based on intention to treat, the small statistically significant effect reported here represents only a small portion of the benefit that patients and the mental healthcare delivery system might gain from real-world use of patient-reported feedback and routine outcomes tracking. If individual patient-reported assessments and outcomes tracking are to reach their full potential, however, they must become a part of routine care, rather than a research endeavor. The IVR system used in this study, paired with a secure Internet-based survey administration system, would exploit this potential by ensuring that screening and assessment instruments are universally accessible to patients. The technology is available now to implement these systems; TeleSage and other organizations are working to integrate IVR and Internet technology to pair universal access to outcomes tracking for patients with real-time report delivery for clinicians.

A number of issues must be addressed before clinical feedback can be made useful on a large scale. The clinical-feedback reports are the gateway to useful outcomes tracking, but if practitioners cannot integrate them into clinical practice, outcomes tracking will have little direct effect. These reports must be brief, clear, and designed to help clinicians focus quickly on those problems that are most likely to improve with clinical intervention. They must be disseminated as quickly as possible, so that clinicians have access to this critical information before the patients present for treatment, and they must be delivered in such a way that clinicians do not feel burdened or threatened by them. Here again, the Internet may offer a solution, in the form of a secure Web site for accessing reports. Finally, clinicians must be trained to understand the reports, and to use them appropriately in clinical practice.

Given the potential benefits to patients, further research into the utility of outcomes tracking and clinical feedback is certainly warranted. The significant improvement attributable to the feedback and the relatively low cost of the intervention suggest that, with the implementation of appropriate systems for gathering and disseminating patient responses, outcomes tracking for clinical use could be a valuable component of patient care.

Acknowledgments

Dr. Brodey thanks Gary Tucker, MD, for his advice on this project and his mentorship. Craig Rosen, PhD, assisted with the statistical analysis for this study. MaryAnne M. Gobble, PhD, contributed to the revision and editing of this article.
REFERENCES


