HHS Public Access

Author manuscript

Psychother Res. Author manuscript; available in PMC 2015 November 01.

Published in final edited form as:

Psychother Res. 2015 November ; 25(6): 694–704. doi:10.1080/10503307.2014.994146.

Predicting Outcome of Substance Abuse Treatment in a Feedback Study: Can Recovery Curves be Improved Upon?

Paul Crits-Christoph,
Room 650, 3535 Market St., Philadelphia, PA 19104, Phone: (215) 662-7993

Hannah M. Markell,
6th Floor, 3535 Market St., Philadelphia, PA 19104, Phone: (215) 349-5222

Robert Gallop,
Department of Mathematics, Applied Statistics Program, West Chester University, 323B Anderson Hall, West Chester, PA 19383, Phone: 610-436-2419

Mary Beth Connolly Gibbons,
Room 649, 3535 Market St., Philadelphia, PA 19104, Phone: (215) 662-3488

Bridget McClure, and
Joint Clinical Trials Office, Weill Cornell Medical College/New York-Presbyterian, 1300 York Avenue, Box 305, New York, NY 10065, Phone: (212) 746-5578

John Rotrosen
423 East 23 Street, VA New York Harbor Healthcare System Floor 17 Room 17014W, Veterans Administration, New York, NY 10010, Phone: 646-754-4763

Paul Crits-Christoph: crits@mail.med.upenn.edu; Hannah M. Markell: hmarkell@mail.med.upenn.edu; Robert Gallop: rgallop@wcupa.edu; Mary Beth Connolly Gibbons: gibbonsm@mail.med.upenn.edu; Bridget McClure: bem254@gmail.com; John Rotrosen: John.Rotrosen@nyumc.org

Abstract

Objective—The goal of the study was to evaluate whether enhanced normative feedback recovery curves are needed for treatment of substance use problems.

Method—Patient predictors of outcome were examined using data from 4 substance abuse treatment clinics.

Results—Baseline severity of symptoms/functioning, employment, and craving were found to be associated with rate of change in symptoms/functioning. Several other variables were associated with rate of change in alcohol use, although in the opposite direction than found in efficacy trials.

Conclusions—The results point to the complexity of designing feedback systems using normative recovery curves for those with substance use problems, and highlight the important differences between real-world treatment of those with substance use problems compared to data from efficacy trials.

Correspondence concerning this article should be addressed to Paul Crits-Christoph, Room 650, 3535 Market Street, Philadelphia PA, 19104. crits@mail.med.upenn.edu.

Keywords

substance abuse; alcohol; feedback; Outcome Questionnaire-45; recovery curve

Substantial evidence has accumulated that providing feedback on patient progress to clinicians can enhance treatment outcomes (Shimokawa, Lambert, & Smart, 2010). In the Lambert et al. (2001, 2001a, 2001b, 2003) feedback system in which clinicians are provided with data based on patient self-assessments using the Outcome Questionnaire-45 (OQ-45) (Lambert, Lunnen, Umphress, Hansen, & Burlingame, 1994), feedback reports incorporate warnings for patients who are not demonstrating expected treatment responses. Expected treatment response is based on normative "recovery curves" derived from large naturalistic treatment samples in which expected rate of change over time is calculated for patients at various levels of initial distress/functioning (Finch, Lambert, & Schaalje, 2001; Lambert et al., 2002). In the OQ-45 feedback studies, patients not progressing as expected are termed *off-track*. Patients who are progressing as expected are referred to as *on-track*.

In an effort to define normative recovery curves for outpatient psychotherapy using an outcome measure similar to the OQ-45, Lueger et al. (2001) created a prediction model based on seven intake variables including subjective well-being, symptoms, life functioning, clinician-rated severity, chronicity, previous treatment, and treatment expectations. However, no information was provided regarding the strength of this predictive model and the amount of variance accounted for by the selected variables. Another approach to feedback is based on a model predicting good/bad outcomes used multiple variables (Kordy, Hannover, & Richard, 2001). For those predicted to achieve poor outcome ("signal cases"), 30% actually achieved poor outcome and 10% achieved relatively better outcome. The authors conclude that "The predictive value of the whole model is not overwhelming" (Kordy et al., 2001, p. 180). When designing the OQ-45 feedback system, recovery curves for patients at different initial levels of the OQ-45 total score were calculated because initial level was a strong predictor of rate of improvement over time. The exact predictive model that is used for the OQ-45 feedback system has not been published, but Lambert et al. (2002) report an R^2 of .17 for a model that contains both initial level and change after one session of therapy predicting final outcome status. For standard outpatient samples, typically mostly comprised of patients with complaints centering on depression, anxiety, and/or interpersonal difficulties, the use of initial level to define normative recovery curves has worked exceptional well, as evidenced by results of feedback studies showing statistically significant and clinically meaningful benefits from feedback based on deviation from such normative recovery curves (Shimokawa, Lambert, & Smart, 2010). Nevertheless, the clinical usefulness of a feedback system might be enhanced by using normative recovery curves based on greater precision of predictive models for change over time.

In the treatment of substance use problems in particular, there are a couple of reasons to believe that other predictor variables, besides initial OQ level, might be relevant to expected rate of improvement. For one, many patients beginning a course of substance abuse treatment are not at their peak level of distress. Patients who arrive for substance abuse treatment have often been clean from drug or alcohol for some period of time before

beginning treatment. Whether treatment is mandated following a driving under the influence or driving while intoxicated conviction, is the result of pressure from family/friends to reduce/stop using drug or alcohol, or is simply a personal decision that it is time to address alcohol/drug use, stopping use in the weeks before an initial treatment session is common. For example, in a recent 4-site substance abuse feedback intervention study, more than 75% of patients reported no days of drug or alcohol use in the past seven days at baseline (Crits-Christoph et al., 2012). Similarly, in the Crits-Christoph et al. (2010) study involving 20 substance abuse treatment clinics, 77% of those who began treatment reported no alcohol use, and 83% reported no drug use, in the past week. When a patient is currently using, the focus of treatment is typically to stop/reduce use of drugs/alcohol. After usage has stopped, treatment often enters a phase of relapse prevention. Because many patients are "clean" from drugs/alcohol at baseline, the goal of treatment for such patients would typically move directly to prevent relapse rather than to focus on stopping current use or reduce immediate distress associated with heavy use. There might be other baseline variables associated with the course of improvement when the focus on relapse prevention rather than stopping current use.

A second reason to expect that other variables besides initial level of use might be important to consider in understanding the typical course of improvement for those with substance use disorders is that a variety of studies have successfully identified predictors of outcome of substance use treatment. In a review of the literature on predictors of substance use outcomes, Haaga, Hall, and Haas (2005) identified three consistent findings in the literature: (a) patients who are employed have better outcomes than those who are unemployed, (b) the presence of comorbid antisocial personality disorder is associated with poorer outcomes, and (c) psychiatric severity is associated with worse outcomes. In a comprehensive study of predictors of the outcome of treatment for cocaine dependence, four baseline variables were identified as significant predictors of outcome (Crits-Christoph et al., 2007). These were: craving, acuity of biomedical problems, expectations for improvement, and belief in the drug counseling recovery (primarily 12-step) philosophy. The joint effect of these predictors was clinically meaningful: for patients with none of the four positive prognostic attributes, the probability of achieving two or more months of consecutive abstinence was 11%; for those with three or more of the prognostic attributes, the probability of achieving two or more months of consecutive abstinence was 62%. Although the Crits-Christoph et al. (2007) predictive results were obtained in the context of a randomized efficacy trial for cocaine dependence, the robustness of the findings across a range of psychosocial treatments (individual and group drug counseling; cognitive therapy, supportive-expressive psychotherapy) in the trial suggests that such predictors might also apply to naturalistically delivered substance abuse treatment (i.e., drug counseling with a 12-step and/or relapse prevention focus).

The current study, therefore, was designed to examine predictors of the outcome of naturalistic substance use treatment in order to inform the development of recovery curves for feedback reports within a substance using population. Based on the Haaga et al. (2005) review, we examined employment status and psychiatric severity (as captured by OQ-45 total scores) as predictors. Based on the Crits-Christoph et al. (2007) study, we examined craving, acuity of biomedical problems, belief in the drug counseling/recovery philosophy,

and expectations for improvement as predictors of change. Given the emphasis on initial severity for defining OQ-45 recovery curves, we also examined baseline levels of drug/alcohol use as predictors for comparison purposes. These analyses were conducted within the context of an OQ-45 feedback study conducted with patients and clinicians in real-world substance abuse treatment facilities. The impact of these predictor variables on rate of change in OQ-45 scores over the course of 12 treatment sessions was investigated for the group of patients whose clinicians did not receive feedback. Because alcohol and drug use outcomes are especially important among those seeking help with substance use problems, we also investigated prediction of these outcomes.

We hypothesized that each of the predictor variables would be associated with rate of change in OQ-45 scores, alcohol use, and drug use. Although normative recovery curves should be evaluated within a non-feedback condition, we also examined the same predictor variables within the context of a feedback condition to explore any potential impact of feedback on the predictor-outcome relations.

METHODS

The current study uses data from a previously reported feedback study (Crits-Christoph et al., 2012) that used a modified version of the OQ-45 feedback system in three real-world clinics that treat patients with substance use problems. There were two phases in the intervention study. In Phase I, patients exclusively were assessed via the modified OQ-45 instrument, administered at the beginning of treatment sessions. In Phase II, different patients at the same clinics completed the modified OQ-45 at their treatment sessions, and their counselors immediately received feedback reports based upon the patient's OQ-45 assessment. Additionally, if a patient went off-track (according to the OQ-45), Phase II counselors were granted access to clinical support tools in addition to feedback. Affiliated university and clinic Institutional Review Boards approved the study protocol.

Participants

Clinics—Four clinics participated in the study, two in New York City, Philadelphia, and Salt Lake City. Each was a community-based outpatient substance abuse treatment clinic (non-methadone maintenance).

Clinicians—Clinician eligibility requirements at the participating clinics were as follows: clinicians needed to be providing individual counseling to adult (age 18 or older) clients, work a minimum of 20% of time at the given clinic, and provide their written informed consent.

Patients—Eligible patients (aged 18 or above) presented at the clinic in search of a new course of individual outpatient treatment and were enrolled within two weeks of their intake session. Additionally, patients could only participate if their individual counseling appointments were scheduled at a minimum of twice monthly. All participating patients gave their written informed consent.

Assessments

Modified Outcome Questionnaire-45 (OQ-45)—The Outcome Questionnaire (OQ-45) was used to track patient progress in this study. The OQ-45 is a 45-item self-report scale which was created specifically for tracking and assessing patient outcomes in a therapeutic setting. The OQ-45 is scored on a 5-point scale (0=*never*, 1=*rarely*, 2=*sometimes*, 3=*frequently*, 4=*almost always*), yielding scores ranging from a possible 0 to 180. Higher scores on the OQ-45 are indicative of greater levels of symptoms and/or poorer functioning. The OQ-45 was modified in one key way for the purposes of the current study, which entailed adding two items which measure the number of days in the past week that the patient used (1) alcohol and (2) drugs. In this study, the total score and the individual alcohol and drug use items were used as outcome measures.

A well-established instrument, the OQ-45 has been validated across a wide array of healthy and clinical populations. The OQ-45's internal consistency reliability (Cronbach's alpha), reported by Lambert et al. (2004) is .93 with a 3-week test–retest reliability value of .84 for the OQ-45 total score. The concurrent validity of the OQ-45 total score has been investigated via correlations with the Symptom Checklist-90 (SCL-90; Derogatis & Cleary, 1977), Beck Depression Inventory (BDI; Beck, Steer, & Garbin, 1988), Zung Depression Scale (Zung, Richards, & Short, 1965), and the State-Trait Anxiety Inventory (STAI; Spielberger, 1983). Each of these concurrent validity correlations were significant at the .01 level with *r*'s ranging from .50 to .85 (Lambert et al., 2004). Perhaps most importantly, the OQ-45 has demonstrated sensitivity to the effects of treatment on patients' functioning (Vermeersch, Lambert, & Burlingame, 2000; Vermeersch et al., 2004).

In both Phases (I and II) the OQ-45 was administered right before every treatment session. Patients could have up to a maximum of 12 post-baseline treatment sessions or six months (whichever came first). Electronically captured data was sent (via encrypted wireless connection) to a computer where it was stored in a database. During Phase II, reports were available to be viewed by clinicians immediately after the patient completed measures via a handheld computer or computer kiosk. The reports were viewed electronically, via a secure, password-protected computer system. Feedback reports generated from OQ-45 scores were the same as reports described in previous studies (Slade, Lambert, Harmon, Smart, & Bailey, 2008), except that patient scores on the alcohol and drug use questions were also included in reports and tracked over time to make changes in drug or alcohol use apparent to the counselor.

During Phase II, when the OQ-45 identified certain patients as being off-track based upon their scores, therapists were given the chance to employ clinical support tools which provided ideas for ways to improve treatment. Upon their identification of being off-track, patients were asked to fill out the Assessment for Signal Clients (ASC; Lambert, et al., 2007), a 40-item self-report measure that is intended to evaluate the severity of certain problems which could be hindering progress in treatment– specifically, problems with therapeutic alliance, social support systems, motivation, and stressful life events.

Addiction Recovery Scale—The Addiction Recovery Scale (ARS) is a self-report scale that was developed for the NIDA Cocaine Collaborative Treatment Study (CCTS). Certain

items on the scale reflect the drug counseling/12-step philosophical system while other items assess involvement in the suggested behaviors. The total score for the ARS is calculated by summing the items (each rated on a 1 to 5 scale). A high score indicates greater endorsement of the philosophy and behaviors encouraged in 12-step oriented drug counseling. The original scale had 40 items. A 10-item version of the scale, used in the current project, was created by selecting the items with the highest item-total correlations (two of the 10 items are reverse scored). Internal consistency (Cronbach's alpha) for the 10-item scale using baseline data in the current project was .75. Baseline ARS total scores have been found to predict sustained abstinence (Crits-Christoph et al., 2007), and change in the ARS total score has been found to be a partial statistical mediator of drug use outcomes of individual (plus group) drug counseling (Crits-Christoph et al., 2003).

Expectations for Improvement—We used one item that directly measured expectations for improvement, "Overall, how much improvement do you expect to experience as a result of treatment here?" This item was rated on a 1 ("I expect to feel worse") to 7 ("I expect to feel much better") scale. Although this was only a single item assessment of expectations for improvement, we used this item as a predictor because a meta-analysis found that ratings of expectations for improvement were consistently related to the outcome of psychosocial treatment across 46 studies, with many of the studies employing a single item scale (Constantino, Arnkoff, Glass, Ametrano, & Smith, 2011).

Craving—Craving for drug/alcohol was assessed each week with a three-item adapted from the Cocaine Craving Scale (Weiss et al., 1997; Weiss et al., 2003) used in the NIDA CCTS. The three items were: (1) "Please rate how strong your desire was to use drugs or alcohol during the last 24 hours," (2) "Please imagine yourself in the environment in which you previously used drugs and/or alcohol. If you were in this environment today, what is the likelihood that you would use drugs or alcohol?", and (3) "Please rate how strong your urges are for drugs or alcohol when something in the environment reminds you of it." Response options ranged from 0 for "no desire/likelihood of use" to 9 for "strong desire/likelihood of use," and the final score was the sum of the three items. Internal consistency of this scale at baseline was .74 in the current study.

Acuity of Biomedical Problems—This scale was taken from the Recovery Attitude and Treatment Evaluator Questionnaire (Mee Lee, Hoffmann, & Smith 1992). The three items of this scale include: 1) "Do you have any physical or medical problems whatsoever?", 2) "Do your physical problems require you to see a health care professional such as a doctor, nurse or rehabilitation specialist?", and 3) "Do your physical problems interfere with your treatment or recovery efforts from drugs/alcohol?" Each item is rated on a scale of 1 (e.g. "None- no need for medical visits") to 4 (e.g. "Very Much- Hospital care is needed"). The scale score is an average of the three items. Internal consistency of this scale at baseline was .74 in the current study.

Procedures

Counselors were given an orientation to the study goals and procedures and trained to interpret feedback reports. If a patient was identified as off-track based upon his rate of

improvement indicated by OQ-45 scores, the ASC was administered via handheld computers or computer kiosks. Once electronically captured, the data was sent to and stored in a computer database via an encrypted wireless connection. In Phase II, the ASC was only completed once—the first time the patient went off-track. Further details of the original feedback study are provided in Crits-Christoph et al. (2012).

Statistical Analysis

Primary outcome analyses included all patients who had at least a baseline and one treatment session's OQ-45 assessment in the Phase I (no-feedback) condition. Because feedback during Phase II influenced outcomes (Crits-Christoph et al., 2012), it was thought that it was more appropriate to primarily examine the relations of predictors to outcome in a treatment-as-usual cohort without feedback, as was done in the original OQ-45 recovery curve study (Lambert et al., 2002), though we also conducted exploratory analyses examining whether predictions varied for no feedback vs. feedback conditions.

The predictor analysis involved a longitudinal analysis of all OQ-45 data collected from session one through session 12. The longitudinal analysis used a random effects multilevel regression model (Goldstein, 1995) with a random slope and random intercept for each subject. Due to non-linear change of individual patient profiles over time, a logarithmic transformation of time was performed. Separate analyses were initially conducted for each predictor variable. The following terms were included in the model: the predictor variable, time, baseline OQ-45 total score, and predictor by time interaction. The latter term was our focus to examine the relation of the predictor to rate of change in OQ-45 scores over time. To test for differential prediction in the feedback and no-feedback conditions, data from both Phase I (no-feedback) and Phase II (feedback) was used, and the weighted averages of subject-specific slope estimates were compared between the feedback and no-feedback groups by including a Predictor by Time by Phase interaction (in addition to the terms listed above, plus Phase and Phase by Time terms).

Drug and alcohol use items of the OQ-45 were analyzed in the same way as OQ-45 total scores. We used PROC MIXED in SAS (Littell, Milliken, Stroup, Wolfinger, & Schbenberger,2006) to conduct the random effects hierarchical regression analyses.

Effect sizes (converted to r), derived from the F-test for mixed effects model, were calculated r = square root of [F /(F + DF)], where F is the F-test statistic for the effect of interest in the multilevel design (Rosenthal & Rosnow, 1991; Verbeke & Molenberg, 2000). We used likelihood-ratio testing to assess the simultaneous effect of all predictors on change, compared to simple model with no predictors. The likelihood ratio test uses a chi-square distribution based on the difference of the log-likelihood functions, referred to as the deviance, with the degrees of freedom equal to the difference in number of fixed pararmeters estimated between the two nested models (Neter, Wasserman, & Kutner, 1989). An effect size (converted to $\eta 2$) was calculated from the likelihood ratio test.

RESULTS

Baseline Characteristics of Sample

Across all three sites, a total of 391 patients attended an intake session, and 304 patients had both intake data and a minimum of one post-baseline assessment in either Phase I or Phase II of the investigation (142 in Pennsylvania; 107 in Utah; 2 in one NY clinic, 53 patients in a second NY clinic). Of these 304 individuals with at least one post-baseline assessment, 165 were from Phase I and 139 were from Phase II.

The 87 patients who attended intake but had no treatment sessions were largely comparable to the 304 patients who had at least one treatment session. Results of chi-square or *t*-test analyses comparing these two groups of patients revealed no significant differences on gender, education, Hispanic ethnicity, marital status, age, baseline alcohol use, or baseline OQ-45. However, there was evidence that those who attended intake only, compared to those who attended at least one treatment session, were more likely to be Black compared to White or Other race ($\chi 2$ [2] = 17.6, $p < .001$), were less likely to be employed ($\chi 2$ [1] = 4.5, $p = .043$; 8.0% of those who attended intake only were employed; 17.3% of those who attended at least one treatment session were employed), and were more likely to report drug use in the past 7 days at intake (intake only group, $M = .97$ days, $SD = 1.9$; at least one treatment session group, $M = .46$ days, $SD = 1.3$; unequal variance *t*-test: t [111] = 2.3, $p = .021$).

Overall, the demographic breakdown of these 304 patients (combining Phase I and Phase II) with baseline assessment and at least one post-baseline assessment was as follows, 55.9% were male and 36.5% were Caucasian. At baseline, 84.2% of patients reported no drug use and 78.6% reported no alcohol use. Alcohol had been a problem for about 11 years for the average patient, while drug use had been a problem for approximately 13 years. No significant differences existed between the Phase I (no-feedback) and Phase II (feedback) samples on any baseline characteristic variables (Table 1).

A total of 38 clinicians across three sites had at least one patient that completed at least one post-baseline assessment. These clinicians were 60.5% female, 28.9% African American, and 7.9% self-identified as Hispanic or Latino. On average, participating clinicians were 44.5 years old. Regarding education, 63.2% of the clinicians had a Master's Degree, 15.8% had a Bachelor's Degree, and 15.8% had a Doctoral Degree. More than half (57.8%) reported having at least five years of experience as a counselor.

At baseline, the OQ-45 total scores correlated .14 ($N = 304$; $p = .008$) with alcohol use and .09 ($N = 304$; $p = .06$) drug use. Alcohol use and drug use were correlated .42 ($N = 304$; $p < .001$) at baseline.

Prediction of Rate of Change in OQ-45

Shown in Table 2 are parameter estimates and the statistical significance of predictor by time interactions (examined individually), representing the prediction of slope over time from session one to session 12, for patients in the no-feedback condition. Also shown are effect sizes expressed as correlations (r).

As expected, baseline level on the OQ-45 total score was a significant predictor of rate of change in the OQ-45 total score from session 1 to session 12 (r = .27, p = .005). However, two other variables, employment and craving at baseline, also significantly predicted rate of change in OQ-45 total scores at nearly the same level of strength (r's of .21 [p = .024], and .26 [p = .009], respectively). Employed patients had a faster rate of change; high craving at baseline was associated with a faster rate of change in OQ-45 total scores. Testing a model in which all three predictor variables were included (main effects plus interactions with time) versus a model without the three interactions with time (only main effects), yielded a highly significant effect for the addition of the three interactions with time (χ^2 [3] = 16.9, p = .00073; effect size: η^2 = .112) However, the individual influence of each was attenuated in the model that contained all three predictors, suggesting that overlap between these variables was partly responsible for their predictive strength (in the model with all three predictors: Baseline OQ-45 total score by Time: r = 18, p = .064; Employment by Time: r = .19, p = .057; Craving by Time: r = .13, p = .205). Of note was that the effect of employment was comparable to that for baseline OQ-45 total score.

Within the feedback condition (Table 3), baseline OQ-45 total score also significantly predicted rate of change in OQ-45 total scores, though the r was slightly lower (r = .24, p = .015). Restricting the sample in the feedback condition to only those who went off-track also revealed a significant relationship between baseline OQ-45 scores and rate of change in the OQ-45 total score (r = .34, p = .03, N = 54). Employment and craving, however, no longer significantly predicted rate of change in OQ-45 scores in the feedback condition (r's of .01 and .16, respectively). To test for the difference in these predictive relations in no-feedback vs. feedback, we examined statistical models that included a Predictor by Time by Phase interaction. Recognizing that statistical power was limited for testing 3-way interactions, we note that no significant results were evident (for Employment, F [1, 256] = 2.3, p = .13; for Craving, F [1, 248] = 0.1, p = .80).

Prediction of Rate of Change in Alcohol and Drug Use

In the no-feedback condition, baseline OQ-45 total score was not associated significantly with rate of change in alcohol use or drug use (Table 2). Rate of change in alcohol use, however, was predicted by several other variables in the no-feedback condition. Baseline level of alcohol use strongly predicted rate of change in alcohol use (r = .49, p < .001). In addition, craving (r = .31, p < .001), acuity of biomedical problems (r = .16, p = .045), and the Addiction Recovery scale (r = .19, p = .011) all significantly predicted rate of change in alcohol use. In addition, a marginally significant (r = .15, p = .052) effect was evident for employment. Relatively greater rate of change in alcohol use was associated with relatively high craving, low acuity of biomedical problems, less belief in the drug counseling/12-step philosophy of recovery, and being employed. In a model that included all of these predictors, baseline alcohol use (r = .42, p < .001) and craving (r = .19, p = .015) remained significant, while the effects of acuity of biomedical problems (r = .09, p = .255) and the Addiction Recovery Scale (r = .12, p = .131) were attenuated.

No significant predictor relations were apparent for rate of change in drug use in the no-feedback group, though a marginally significant (r = .19, p = .052) effect was evident for baseline level of drug use as a predictor.

Within the feedback group, craving (r = .22, p = .028), the Addiction Recovery Scale (r = .27, p = .005) and baseline level of alcohol use (r = .33, p < .001) all continued to be significant predictors of rate of change in alcohol use. While not significant, the relation of Acuity of Biomedical Problems to rate of change in alcohol use was of similar magnitude (r = .15) to that found in the no-feedback group. Craving (r = .20, p = .039) and the Addiction Recovery Scale (r = .23, p = .014) also predicted rate of change in drug use, as did baseline level of drug use (r = .22, p = .031). No significant interactions of predictors by phase (feedback vs. no-feedback) were apparent for rate of change in alcohol or drug use.

DISCUSSION

The results of the current study have several implications for the use feedback systems in a substance using population. First, we found that baseline levels on the OQ-45 total score were associated with rates of change in OQ-45 total scores. This finding appears to validate, within a substance using population, the use of normative recovery curves for the OQ-45 that are based on initial level of the OQ-45 total score. However, initial level on the OQ-45 was not the only significant predictor of rate of change in the OQ-45. Two other variables, employment and craving, also predicted rate of change in the OQ-45. These findings therefore potentially suggest that more accurate normative OQ-45 recovery curves for a substance using population could be generated by incorporating multiple variables into models for recovery curves. Although the current OQ-45 feedback system produced positive results within a substance using population (Crits-Christoph et al., 2012), it is possible that more accurate normative recovery curves would allow for a feedback system that targets clinicians' attention more successfully to those patients who are deviating from the rate of change that one would expect for similar patients.

Our finding for employment is consistent with the Haaga et al. (2005) review that found that employment was a consistent predictor of the outcome of substance abuse treatments. Based on the consistency of this finding, it would seem useful for feedback systems to alert clinicians to whether a particular patient is improving as expected, or not, relative to other patients with the same employment status. However, the finding for craving seems to be inconsistent with the Crits-Christoph et al. (2007) study conducted using a cocaine dependent sample. In the current study, high levels of craving were associated with a faster rate of change. In the Crits-Christoph et al. (2007) study, high rates of craving at baseline was associated with less duration of sustained abstinence over the course of treatment. Two factors are important to consider in understanding these discrepant findings in regard to craving. One is that, given that in the current sample about 80% of patients were not using drugs or alcohol in the past week at baseline, craving was relatively low for the sample as a whole. In fact, 21.3% of patients reported no craving or desire to use at all (score of 0 summing across three zero to nine craving items) in the current sample. The relatively low levels of craving may have created a "floor" effect where those at the low end had little room to improve on the OQ-45 and no room to improve on the alcohol and drug use

outcomes (zero use). Thus, high levels of craving could not be associated with relatively poor outcome in the current context. Secondly, it may be that craving is a stronger predictor of poor outcome within a cocaine dependent (or opioid dependent) sample compared to a mixed substance use sample (many with alcohol as the primary problem) that may be more characteristic of patients seeking treatment in some real-world substance abuse treatment facilities.

Also of interest in the current study was that craving, acuity of biomedical problems, and belief in the drug counseling/12-step philosophy were all predictive of rate of change in alcohol use. Each of these variables, however, predicted outcome in the opposite direction as was found in the Crits-Christoph et al. (2007) cocaine study. In addition, higher baseline levels of alcohol use were associated with a faster rate of change in alcohol use in the current study. In Project MATCH, a higher level of alcohol involvement prior to treatment was related to more drinks per day during follow-up (Project Match Research Group, 1997). As with craving predicting rate of change in the OQ-45, the current results for predictors of change in alcohol use are likely a function of the low level of baseline alcohol use for many patients, and the impact of this “floor” effect on predictive relationships. In addition, although those with relatively higher alcohol use at baseline may change more in the short-run, such patients may also be more vulnerable to relapse at follow-up. Thus, predictors of short-term and long-term outcomes in the treatment of substance use disorders may differ.

The nature of the results of the current study, with multiple variables statistically significant in opposite directions than found previously, highlight the important distinction between the population and context of randomized clinical trials compared to naturalistic real-world clinics in the treatment of individuals with substance use problems. For randomized clinical trial efficacy studies, patients are pre-selected to have a pre-determined degree of severity of the target problem at baseline so that change can be detected on that problem over the course of treatment. This typically translates into selecting patients who qualify for a DSM diagnosis of abuse/dependence and/or have used drugs/alcohol recently. As we have seen in the current study and other studies, many patients in real-world substance abuse treatment clinics have not used drug/alcohol recently. This may be a key factor in influencing the role of predictor variables.

Beyond the problem of low levels of substance use at baseline, patients seeking help at substance abuse treatment programs are heterogeneous on a variety of meaningful variables. Some patients are mandated to treatment, and the incentives connected to this process likely override the influence of other variables that normally would be associated with treatment outcome. In addition, individuals seek help in community clinics for problems with a wide range of drugs of abuse, including alcohol, opioids (both prescription medications and heroin), cocaine, methamphetamine, marijuana, club drugs, hallucinogens, inhalants. The course of recovery from these different drugs, as well as polydrug abuse, may be somewhat different. Course of recovery is also likely to vary as a function of whether a patient is currently using or has been abstinent for a period of time. In addition, antisocial personality disorder has been consistently linked to poor outcome of substance abuse treatment (Haaga et al., 2005). Genetic variability may also play a role in terms of responsiveness to treatment for substance use problems (e.g., Anton et al., 2008). Because of the potential influence of

all of these variables, identifying “normative recovery curves” in the context of such patient heterogeneity for those seeking help at substance abuse treatment programs would seem to be a major challenge.

An alternative to a single set of recovery curves based on multiple variables that would apply to all patients seeking treatment at substance abuse programs would be to develop recovery curves for separate subpopulations (e.g., young marijuana users; current heavy alcohol users, etc.). Obtaining adequate (large sample) databases for each subgroup, and then programming a feedback system to incorporate a diverse set of recovery curves for identifying off-track patients, is not likely to occur in the short-run. Even if successful, the resultant system, tied to various recovery curves depending on a large number of baseline patient characteristics, might be so complicated that clinical use is limited.

Ideally, a study would be done to test whether feedback based on enhanced recovery curves leads to clinically significant improvement compared to feedback based on recovery curves developed only based on one predictor (i.e., initial OQ total score level). Although such finely tuned recovery curves may not be feasible, our results indicate that some adaptation of recovery curves based on certain findings (i.e., employment status) may improve the predictive validity of current feedback systems. It is likely that the clinical usefulness of more refined recovery curves is dependent on the size of the predictor-outcome relations on which such recovery curves are based; however, it is not known how high an effect is needed to have a meaningful impact when more refined recovery curves are used as the basis for feedback reports. Though the strength of individual predictors in the current study, as typically in other studies, is relatively small, if the joint effect of multiple predictors reaches a moderate or larger effect, it may well be worth generating recovery curves based on these predictors.

In addition to their potential role as part of the generation of recovery curves, it may also be of interest to examine the effects of feedback within subgroups defined by certain clinically relevant baseline variables. It may be, for example, that feedback has little impact among patients who have a long history of substance use. Such possible moderators of feedback effects should be tested in future studies with large samples sizes so that adequate statistical power exists for examining feedback effects within subgroups.

A number of limitations to the current study need mention. One limitation of this study is the lack of diagnostic information on patients. This makes it difficult to fully characterize the sample. Another limitation is that only certain potential predictor variables were available for analysis. A wider range of variables needs to be examined in future research to more thoroughly investigate the potential predictors of rate of change in a naturalistic sample of patients seeking help at substance abuse treatment programs. Although data was collected at four clinics in separate areas of the United States, the generalizability of the current results to other locations is not known. Another limitation was that the sample that attended intake only was different on three variables from the sample of patients who attended at least one treatment session. Thus, the generalizability of the results to all patients who initially present to a substance abuse treatment facility is not known. However, it may be that it is more

appropriate to generalize data on recovery curves only to the subset of individuals who actually receive any treatment.

Despite these limitations, the results support that recovery curves based on initial level are valid but might be improved based on other important baseline variables like employment. Substantial further research would be needed to fine-tune the accuracy of recovery curves for specific subgroups; though whether this would be clinically useful remains to be determined.

Acknowledgments

Paul Crits-Christoph, Hannah M. Markell, and Mary Beth Connolly Gibbons, Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, USA; Robert Gallop; Department of Mathematics, West Chester University, West Chester PA, USA; Bridget McClure, Office of Joint Clinical Trials, Weill Cornell Medical College, NY, USA; and John Rotrosen, Department of Psychiatry, New York University School of Medicine and DVA New York Harbor Healthcare System, New York, NY, USA

This work was supported by National Institute on Drug Abuse grants R01-DA020799 and R01-DA020809.

References

Anton RF, Oroszi G, O'Malley S, Couper D, Swift R, Pettinati H, Goldman D. An evaluation of mu-opioid receptor (OPRM1) as a predictor of naltrexone response in the treatment of alcohol dependence: results from the Combined Pharmacotherapies and Behavioral Interventions for Alcohol Dependence (COMBINE) study. Archives of General Psychiatry. 2008; 65:135–144.10.1001/archpsyc.65.2.135. [PubMed: 18250251]

Beck AT, Steer RA, Garbin MG. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. Clinical Psychology Review. 1988; 8:77–100.10.1016/0272-7358(88)90050-5

Constantino MJ, Arnkoff DB, Glass CR, Ametrano RM, Smith JZ. Expectations. Journal of Clinical Psychology. 2011; 67:184–192.10.1002/jclp.20754 [PubMed: 21128304]

Crits-Christoph P, Connolly Gibbons MB, Barber JP, Hu B, Hearon B, Worley M, Gallop R. Predictors of sustained abstinence during psychosocial treatments for cocaine dependence. Psychotherapy Research. 2007; 17:240–252.10.1080/10503300600818210

Crits-Christoph P, Gibbons MB, Barber JP, Gallop R, Beck AT, Mercer D, Tu X, Thase ME, Weiss RD, Frank A. Mediators of outcome of psychosocial treatments for cocaine dependence. Journal of Consulting and Clinical Psychology. 2003; 71:918–925.10.1037/0022-006X.71.5.918 [PubMed: 14516240]

Crits-Christoph P, Ring-Kurtz S, Hamilton J, Lambert M, Gallop R, McClure B, Rotrosen R. A preliminary study of the effects of individual patient-level feedback in outpatient substance abuse treatment programs. Journal of Substance Abuse Treatment. 2012; 42:301–309.10.1016/j.jsat.2011.09.003 [PubMed: 22036697]

Derogatis LR, Cleary PA. Confirmation of the dimensional structure of the SCL-90: A study in construct validation. Journal of Clinical Psychology. 1977; 33:981–989.10.1053/eujp.2000.0162

Finch AE, Lambert MJ, Schaalje BG. Psychotherapy quality control: The statistical generation of expected recovery curves for integration into an early warning system. Clinical Psychology and Psychotherapy. 2001; 8:231–242.10.1002/cpp.286

Goldstein, H. Multilevel statistical models. 2nd. Edward Arnold; London: 1995.

Haaga, DAF.; Hall, SM.; Haas, A. Participant factors in substance use disorders. In: Castonguay, L.; Beutler, L., editors. Principle of therapeutic change that work. NY: Oxford University Press; 2005.

Harmon SC, Lambert MJ, Smart DM, Hawkins E, Nielsen SL, Slade K, Lutz W. Enhancing outcome for potential treatment failures: Therapist-client feedback and clinical support tools. Psychotherapy Research. 2007; 17:379–392.10.1080/10503300600702331

Kordy H, Hannover W, Richard M. Computer assisted feedback driven active quality management for psychotherapy provision: The Stuttgart – Heidelberg model. Journal of Consulting and Clinical Psychology. 2001; 69:173–183.10.1037/0022-006X.69.2.173 [PubMed: 11393595]

Lambert, MJ.; Bailey, R.; Kimball, K.; Shimokawa, K.; Harmon, SC.; Slade, K. Clinical Support Tool Manual-Brief Version-40. Salt Lake City: OQMeasures; 2007.

Lambert M, Hansen N, Finch A. Patient-focused research: Using patient outcome data to enhance treatment effects. Journal of Consulting and Clinical Psychology. 2001; 69:159–172.10.1037/0022-006X.69.2.159 [PubMed: 11393594]

Lambert MJ, Hawkins EJ. Using information about patient progress in supervision: Are outcomes enhanced? Australian Psychologist. 2001; 36:131–138.10.1080/00050060108259645

Lambert, MJ.; Morton, JJ.; Hatfield, D.; Harmon, C.; Hamilton, S.; Reid, RC.; Burlingame, GB. Administration and scoring manual for the Outcome Questionnaire_45. Orem: UT American Professional Credentialing Services; 2004.

Lambert MJ, Whipple JL, Hawkins EJ, Vermeersch DA, Nielsen SL, Smart DW. Is it time for clinicians to routinely track patient outcome? A meta-analysis. Clinical Psychology. 2003; 10:288–301.10.1093/clipsy.bpg025

Lambert MJ, Whipple JL, Smart DW, Vermeersch DA, Nielsen SL, Hawkins EJ. The effects of providing therapists with feedback on patient progress during psychotherapy: Are outcomes enhanced? Psychotherapy Research. 2001; 11:49–68.10.1093/ptr/11.1.49 [PubMed: 25849877]

Lambert MJ, Whipple JL, Bishop MJ, Vermeersch DA, Gray GV, Finch AE. Comparison of empirically derived and rationally derived methods for identifying clients at risk for treatment failure. Clinical Psychology and Psychotherapy. 2002; 9:149–164.10.1002/cpp.333

Littell, RC.; Milliken, GA.; Stroup, WW.; Wolfinger, RD.; Schabenberger, O. SAS System for Mixed Models. Second. Cary, N.C: SAS Institute; 2006.

Lueger RJ, Howard KI, Martinovich Z, Lutz W, Anderson EE, Grissom G. Assessing treatment progress with individualized models of expected response. Journal of Consulting and Clinical Psychology. 2001; 69:150–158.10.1037/0022-006X.69.2.150 [PubMed: 11393593]

Mee-Lee, D.; Hoffmann, NG.; Smith, MB. Recovery Attitude and Treatment Evaluator (RAATE) Manual. St. Paul, MN: CATOR/New Standards; 1992.

Neter, J.; Wasserman, W.; Kutner, MH. Applied Linear Regression Models. 2nd. Burr Ridge, IL: Irwin Inc.; 1989. See p. 604

Project MATCH Research Group. Matching alcoholism treatments to client heterogeneity: Project MATCH posttreatment drinking outcomes. Journal of Studies on Alcohol. 1997; 58:7–29. Retrieved from: http://www.jsad.com/jsad/downloadarticle/Matching_Alcoholism_Treatments_to_Client_Heterogeneity_Project_MATCH_Postt/440.pdf. [PubMed: 8979210]

Rosenthal, R.; Rosnow, RL. Essentials of behavioral research: Methods and data analysis. 2nd. New York: McGraw-Hill; 1991.

Shimokawa K, Lambert MJ, Smart DW. Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic and mega-analytic review of a psychotherapy quality assurance system. Journal of Consulting and Clinical Psychology. 2010; 78:298–311.10.1037/a0019247 [PubMed: 20515206]

Slade K, Lambert MJ, Harmon SC, Smart DW, Bailey R. Improving psychotherapy outcome: The use of immediate electronic feedback and revised clinical support tools. Clinical Psychology & Psychotherapy. 2008; 15:287–303.10.1002/cpp.594 [PubMed: 19115449]

Spielberger, CD. Manual for the State-Trait Anxiety Inventory (STAI). Palo Alto, CA: Consulting Psychologists Press; 1983.

Verbeke, G.; Molenberg, G. Linear Mixed Models for Longitudinal Data. New York: Springer; 2000.

Vermeersch DA, Lambert MJ, Burlingame GM. Outcome questionnaire: Item sensitivity to change. Journal of Personality Assessment. 2000; 74:242–261.10.1037/0022-0167.51.1.38 [PubMed: 10879354]

Vermeersch DA, Whipple JL, Lambert MJ, Hawkins EJ, Burchfield CM, Okiishi JC. Outcome Questionnaire: Is It Sensitive to Changes in Counseling Center Clients? Journal of Counseling Psychology. 2004; 51:38–49.10.1037/0022-0167.51.1.38

Weiss RD, Griffin ML, Hufford C, Muenz LR, Najavits LM, Jansson SB, Thompson HJ. Early prediction of initiation of abstinence from cocaine. The American Journal on Addictions. 1997; 6:224–231.10.1111/j.1521-0391.1997.tb00401.x [PubMed: 9256988]

Weiss RD, Griffin ML, Mazurick C, Berkman B, Gastfriend DR, Frank A, Moras K. The relationship between cocaine craving, psychosocial treatment, and subsequent cocaine use. American Journal of Psychiatry. 2003; 160:1320–1325.10.1176/appi.ajp.160.7.1320 [PubMed: 12832248]

Zung WW, Richards CB, Short MJ. Self-Rating Depression Scale in an outpatient clinic: Further validation of the SDS. Archives of General Psychiatry. 1965; 13:508–515. Retrieved from http://archpsyc.ama-assn.org/cgi/reprint/13/6/508. [PubMed: 4378854]

Table 1

Characteristics of Patients

Characteristic	No-Feedback (Phase I) *N*=165	Feedback (Phase II) *N*=139	Total *N* =304
Gender (women), *n*/*N* (%)	64/165 (38.8%)	58/139 (41.7%)	122/304 (40.1%)
Age, *M*,*N* (*SD*)	39.3, 163 (10.9)	39.0, 139 (9.9)	38.9, 302 (10.4)
Race, *n*/*N* (%)			
Caucasian	65/163 (39.9 %)	55/138 (39.9%)	120/301 (39.9%)
Black	66/163 (40.5%)	64/138 (46.4%)	130/301 (43.2%)
Other Race	32/163 (19.6%)	19/138 (13.8%)	51/301 (16.9%)
Ethnicity (Latino), n/*N* (%)	25/159 (15.7%)	16/136 (11.8%)	41/295 (13.9%)
Marital Status, n/*N* (%) (Married or Cohabitating)	29/162 (17.9%)	27/135 (20.0%)	56/297 (18.9%)
Years of School, n/ *N* (%) (H.S. Diploma and further)	110/163 (67.5%)	101/136 (74.3%)	211/299 (70.6%)
Employment Status, n/*N* (%) (Employed full or part-time)	27/162 (16.7%)	25/138 (18.1%)	52/300 (17.3%)
Alcohol Use in past week, n/*N* (%) at baseline (1 or more days of use)	30/165 (18.2%)	35/139 (25.2%)	65/304 (21.4%)
Drug Use in past week, n/*N* (%) at baseline (1 or more days of use)	27/165 (16.4%)	21/139(15.1%)	48/304 (15.8%)
Mean OQ45Total Score at baseline, *M*, *N* (*SD*)	70.6, 165 (28.7)	68.6, 139 (28.3)	69.7, 304 (28.5)

Note. Sample sizes reflect small amounts of missing data for some variables.

Table 2

Significance Testing for Predictors of Outcome: No-Feedback Condition

	Rate of Change in OQ-45 Total Score				Rate of Change in Alcohol Use				Rate of Change in Drug Use			
Predictor	**Estimate (se)**	***F (DF)***	***p***	***r***	**Estimate (se)**	***F (DF)***	***p***	***r***	**Estimate (se)**	***F (DF)***	***p***	***r***
Employed by Time	−12.497 (5.470)	5.2 (109)	.024	.21	−.646 (.330)	3.8 (161)	.052	.15	−.087 (.283)	0.1 (96)	.759	.03
Expectations for Improvement by Time	−.752(2.269)	0.1 (110)	.741	.03	−.015 (.134)	0.0 (185)	.908	.00	−.007 (.111)	0.0 (106)	.951	.00
Acuity of Biomedical Problems by Time	−1.884 (3.052)	0.4 (108)	.538	.06	−.371 (.184)	4.1 (165)	.045	.16	.259 (.152)	2.9 (96)	.092	.17
Craving by Time	−.837 (.314)	7.1 (96)	.009	.26	−.074 (.018)	17.0 (163)	<.001	.31	−.018 (.016)	1.2 (90)	.275	.12
Addiction Recovery Scale by Time	3.452 (3.094)	1.2 (100)	.267	.11	.485 (.189)	6.6 (171)	.011	.19	.093 (.158)	0.3 (98)	.559	.06
Baseline OQ-45 Total Score by Time	−.214 (.075)	8.1 (106)	.005	.27	−.008 (.005)	3.2 (177)	.075	.13	−.003 (.004)	0.8 (97)	.387	.09
Baseline Level of Alcohol Use by Time	−2.783 (1.666)	2.8 (87)	.098	.18	−.601 (.092)	42.5 (135)	<.001	.49	−.062 (.084)	0.5 (78)	.466	.08
Baseline Level of Drug Use by Time	−1.295 (1.831)	0.5 (118)	.481	.07	−.201 (.106)	3.6 (195)	.060	.13	−.177 (.090)	3.8 (102)	.052	.19

Note. Separate models were calculated for each predictor. The Predictor by Time term examines the prediction of rate of change over time (slope), using a log transformation of session number. The baseline score on the dependent variable, obtained at intake, is also included in the model. The dependent variables are scores from session 1 (which occurs subsequently to baseline) to session 12 (or earlier endpoint for those who terminate). Alcohol and drug use are days using in the past 7 days.

Table 3

Significance Testing for Predictors of Outcome: Feedback Condition

	Rate of Change in OQ45 Total Score				Rate of Change in Alcohol Use				Rate of Change in Drug Use			
Predictor	**Estimate (se)**	***F (DF)***	***p***	***r***	**Estimate (se)**	***F (DF)***	***p***	***r***	**Estimate (se)**	***F (DF)***	***p***	***r***
Employed	1.073 (8.474)	0.0 (107)	.900	.01	−.218 (369)	0.4 (109)	.556	.06	−.631 (.427)	2.2 (116)	.142	.14
Expectations for Improvement by Time	2.617 (2.471)	1.1 (109)	.292	.10	.021 (.114)	0.2 (125)	.857	.04	−.124 (.129)	0.9 (126)	.336	.01
Acuity of Biomedical Problems by Time	−6.484 (4.579)	2.0 (92)	.160	.15	−.284 (.199)	2.0 (93)	.157	.15	−.299 (.229)	1.7 (91)	.195	.14
Craving by Time	−.692 (.429)	2.6 (95)	.110	.16	−.041 (.019)	5.0 (97)	.028	.22	−.045 (.022)	4.4 (100)	.039	.20
Addiction Recovery Scale by Time	9.350 (4.205)	4.9 (106)	.028	.21	.510 (178)	8.2 (106)	.005	.27	.531 (.212)	6.2 (114)	.014	.23
Baseline OQ-45 Total Score by Time	−.279 (.113)	6.1 (96)	.015	.24	−.006 (.005)	1.6 (103)	.203	.13	−.001 (.006)	0.0 (105)	.841	.02
Baseline Level of Alcohol Use by Time	.490 (2.446)	0.0 (127)	.842	.02	−.375 (.101)	13.8 (104)	.000	.33	−.113 (.128)	1.0 (124)	.379	.08
Baseline Level of Drug Use by Time	.165 (2.413)	0.0 (98)	.946	.01	−.206 (.103)	4.0 (88)	.049	.21	−.265 (.121)	4.8 (95)	.031	.22

Note. Separate models were calculated for each predictor. The Predictor by Time term examines the prediction of rate of change over time (slope), using a log transformation of session number. The baseline score on the dependent variable, obtained at intake, is also included in the model. The dependent variables are scores from session 1 (which occurs subsequently to baseline) to session 12 (or earlier endpoint for those who terminate). Alcohol and drug use are days using in the past 7 days.