Predicting Dropout from Inpatient Substance Use Disorder Treatment: A Prospective Validation Study of the OQ-Analyst

Substance Abuse: Research and Treatment
Volume 13: 1–9

DOI: 10.1177/1178221819866181

Hanne H Brorson[1,2], Espen Ajo Arnevik[2] and Kim Rand[3]

[1]Department of Psychology, University of Oslo, Norway, Oslo. [2]Department of Substance use Disorder Treatment, Oslo University Hospital, Oslo. [3]Health Services Research Unit, Akershus University Hospital, Lørenskog, Norway.

ABSTRACT

BACKGROUND AND AIMS: There is an urgent need for tools allowing therapists to identify patients at risk of dropout. The OQ-Analyst, an increasingly popular computer-based system, is used to track patient progress and predict dropout. However, we have been unable to find empirical documentation regarding the ability of OQ-Analyst to predict dropout. The aim of the present study was to perform the first direct test of the ability of the OQ-Analyst to predict dropout.

DESIGN: Patients were consecutively enlisted in a naturalistic, prospective, longitudinal clinical trial. As interventions based on feedback from the OQ-Analyst could alter the outcome and potentially render the prediction wrong, feedback was withheld from patients and therapists.

SETTING: The study was carried out during 2011–2013 in an inpatient substance use disorder clinic in Oslo, Norway.

PARTICIPANTS: Patients aged 18 to 28 years who met criteria for a principal diagnosis of mental or behavioural disorder due to psychoactive substance use (ICD 10; F10.2–F19.2).

MEASUREMENTS: Red signal (predictions of high risk) from the Norwegian version of the OQ-Analyst were compared with dropouts identified using patient medical records as the standard for predictive accuracy.

FINDINGS: A total of 40 patients completed 647 OQ assessments resulting in 46 red signals. There were 27 observed dropouts, only one of which followed after a red signal. Patients indicated by the OQ-Analyst as being at high risk of dropping out were no more likely to do so than those indicated as being at low risk. Random intercept logistic regression predicting dropout from a red signal was statistically nonsignificant. Bayes factor supports no association.

CONCLUSIONS: The study does not support the predictive ability of the OQ-Analyst for the present patient population. In the absence of empirical evidence of predictive ability, it may be better not to assume such ability.

KEYWORDS: feedback technology, progress monitoring, OQ-Analyst, OQ-45, substance use disorder treatment, dropout, attrition, negative treatment outcome, prediction

RECEIVED: June 29, 2019. **ACCEPTED:** July 5, 2019.

TYPE: Original Research

FUNDING: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The study was indirectly financed by the Norwegian Research Council as part of the PhD scholarship to the first author. The funding source had no involvement in the study.

DECLARATION OF CONFLICTING INTEREST: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Hanne H Brorson, Department of Psychology, University of Oslo, 0317, Norway. Email: hanne@brsa.no

CLINICALTRAILS.GOV IDENTIFIER: NCT01891045

Introduction

Addiction to drugs and alcohol is a pervasive problem with harmful consequences for the addicted individual and high costs to society.[1–5] In order to counteract addiction, different treatments have been developed, and one of the most consistent factors related to successful outcomes is treatment completion.[6] Unfortunately, it is more common for a patient to drop out of treatment than to complete.[7] The potential to prevent future dropout hinges on the ability to successfully identify individuals at risk of dropping out in time to allow for intervention. To date, dropout research has predominantly focused on pretreatment predictors, emphasising demographics such as age and sex.[8] With few exceptions, this line of research has been unsuccessful,[8–11] and there has been a call to shift focus to treatment process factors and to systems that help therapists monitor and assess individual treatment process.[12–15] Among the reported benefits of employing progress monitoring (PM) systems is reduced dropout[16–21] and several organizations (e.g. American Psychiatric Association) and national guidelines (e.g. USA and Australia) recommend using PM as means of ensuring treatment quality.[22]

There are several PM systems to choose from;[22] one of the most popular and extensively researched systems is the OQ-Analyst.[23–25] The system has yielded impressive results in mental health services with effect sizes of different types of feedback ranging from .28 to .70,[19] and has been accepted into the National Registry of Evidence-Based Programs and Practices[26] as an evidence-based intervention for substance use disorder (SUD) treatment in inpatient and outpatient settings. The OQ-Analyst includes an algorithm allowing prediction of

dropout of treatment.[27] This is stated explicitly in the feedback message to the therapist following a red alert:

> '[. . .] Chances are your client will drop out of treatment prematurely or have a negative treatment outcome. [. . .]'[28]

While the system has a proven track record for predicting negative treatment outcome,[29–34] we have been unable to find empirical support for the claim to predict dropout. The lack of documentation was confirmed by one of the founders of the system (Lambert, 2012, personal communication); the present study was therefore designed to perform the first direct test of the ability of the OQ-Analyst to predict dropout. A frequent phenomenon such as dropout should be more easily predicted than infrequent events such as deterioration;[28] thus the OQ-Analyst would be expected to predict dropout above the accuracy level documented for deterioration. Previous studies investigating prediction of deterioration have concluded with hit rates between 85% and 100%,[29–34] and we therefore hypothesised that the OQ-Analyst would predict dropout at an accuracy of above 85%.

Method

Study design

The study was a naturalistic, prospective, single-centre, longitudinal study designed to investigate the ability of the OQ-Analyst to predict dropout from an inpatient SUD treatment. Unlike the original use of the OQ-Analyst where feedback is available to the patient and therapist shortly after submission, the feedback in the present study was stored away from patients and therapists. As we were investigating the *predictive* ability of the OQ-Analyst, it was pivotal that the therapists did not intervene based on the feedback from the OQ-Analyst as that could alter the outcome and thus render the prediction wrong.

The study was part of the thematic health register Youth Addiction Treatment Evaluation Project (YATEP). YATEP is approved by the Norwegian Data Service for Research, the Norwegian Regional Committees for Medical Research Ethics (REK) and the Oslo University Hospital Data Protection Officer. All procedures for the current study were approved by REK (22 September 2011) (2011/1745) and performed according to REK guidelines and the Helsinki Declaration. The protocol and supporting STARD checklist are available as supporting information; see S1 Protocol, S2 STARD Checklist.

Participants

Data were collected from 1 January 2011 to 31 December 2013 at the Department of Addiction Treatment Youth at Oslo University Hospital, a specialised public hospital in Norway. To ensure the external validity and clinical utility of our findings, study inclusion was designed to follow the intake procedures at the hospital. Eligible patients were those: (1) aged 18 to 28 years; (2) who met criteria for a principal diagnosis of mental or behavioural disorder due to psychoactive substance use (ICD 10; F10.2–F19.2) (3) in need of interdisciplinary services; (4) provided in an inpatient treatment setting. Patients were excluded from the study if they had submitted fewer than two OQ-Analyst assessments (as two assessments is the absolute minimum required to predict dropout) or if they had an insufficient understanding of the Norwegian language to answer a questionnaire without assistance.

Data were collected as part of routine care, but patients were offered the option to refuse participation. Patients providing a written informed consent were asked to fill out the OQ-Analyst questionnaire using an on-site computer once every week, on a fixed day, throughout the course of treatment.

Fifty-four patients (70.1% male, mean age 24 years, SD = 2.42) were assessed for eligibility. One patient refused participation, two were transferred to outpatient treatment, and 11 patients submitted fewer than two OQ forms. The remaining 40 patients were included in the sample and analysed. See Table 1 for sample characteristics and Figure 1 for STARD flow chart of study enrolment.

Test methods

Index test. The OQ-Analyst is a computer-based system used to track patient progress and predict treatment outcome.[24] Data for the OQ-Analyst are typically collected before each treatment session using the Outcome Questionnaire 45 (OQ-45), a 45-item self-report scale assessing three aspects of mental health over the past week: (1) symptoms; (2) relationships; and (3) functioning. After submission, the OQ-Analyst generates a report showing the patient`s session-by-session progress and his or her predicted treatment outcome.

The psychometric properties of the OQ-Analyst have been studied extensively and results show that system is highly reliable, valid and sensitive to changes patients make during treatment.[24,35–37] The system has been validated across a broad range of settings, populations and countries,[26] and the Norwegian version of the *OQ-Analyst* used in this study has been shown to have adequate reliability, internal stability (r = .85 and .93, respectively) and validity with other international instruments.[35]

Reference standard. Treatment personnel at the Department of Addiction Treatment Youth are required to document every dropout incident in the patient's medical record; this was used as the standard for predictive accuracy. Two hospital nurses extracted information about dropout events independently; any disagreement between the two were solved through a discussion with the first author, a licensed clinical psychologist.

Outcome

As there is currently no universally accepted definition of dropout in substance abuse research,[8] we chose to follow the hospital

Table 1. Characteristics of the sample (n = 40).

CHARACTERISTICS	
Sex (%)	
Male	70.1
Female	29.9
Mean age at first admission (years)	24 (SD = 2.42)
Mean years of school	11 (SD = 1.55)
Main ICD 10 substance-related disorder (n (%))	
Opioid	14 (35)
Other stimulants	10 (25)
Cannabinoid	9 (22.5)
Tentative or missing	7 (17.5)
Main comorbidity on axis I and II (n (%))	
Mood disorder	9 (22.5)
Personality disorder	5 (12.5)
ADHD	4 (10)
PTSD	4 (10)
Tentative or missing	18 (45)
Mean length of stay (days)	112.28 (SD = 85.91)
Mean time until first dropout	59.17 (SD = 66.24)
Mean number of dropout	.68 (SD = .94)
Mean baseline OQ-45 score	84 (SD = 22.14)

Abbreviations: ADHD, attention deficit hyperactivity disorder; PTSD, post-traumatic stress disorder.

practice of recording *dropout* whenever a patient discontinued treatment before the treating personnel recommended discharge. This is a common classification of dropout known as discharge against medical advice (AMA).[38] Conceptually, we can find two kinds of patients subsumed under this definition: (1) the patient that permanently leaves treatment; and (2) the patient who returns to treatment. For the latter category, it should be noted that the continuation of treatment after a dropout was decided individually and should not be confused with an 'open-return' policy. A large number of sensitivity analyses using several cut-offs for how long a treatment absence had to last in order to be recorded as a dropout can be found in the 'Results' section.

Predictor

When using the OQ-Analyst as intended, the therapists are alerted to risk of dropout by a red colour signal found in the OQ report. The red signal indicates those patients who are progressing significantly less than expected rates of improvement based upon samples with patients at the same initial level of distress. Accordingly, a red signal was used as predictor of dropout in the current investigation.

Analyses

Power. To determine an appropriate sample size, we used simulation models to create virtual datasets similar in form to those that would be produced by the trial; planned analyses were run on these simulated datasets. This allowed us to estimate power as a function of number of participants, as well as to vary their expected length of stay, rate of dropout, and the expected predictive accuracy of red signal from the OQ-Analyst instrument. The simulation models were based on the following expectations: overall risk of dropout was held constant over time. Reported rates of dropout vary substantially by study population and definition of dropout.[8] For a rough estimate, we calculated a weighted average based on 91 of the studies analysed by Brorson et al (2013)[8] for which single dropout rates were reported, of 48.6%, or roughly half. Assuming a treatment period of approximately three months, or 13 weeks, a fixed weekly risk of 5% corresponds to an overall dropout rate of approximately 50%. The red signal was expected to occur following approximately 10% of all completed OQ questionnaires. Simulated patients' length of stay was modelled to vary depending on modelled dropout. With dropout risk approximately half of the expected frequency of the red signal, a sensitivity of 100% would imply a positive predictive value of 50%. Note that with 10% of all OQ responses yielding a red signal, and a 5% weekly risk of dropout, 50% sensitivity would imply a relative risk of dropout from red signal of 9. We varied the sensitivity of the red signal by increments of 5% from 5% to 95%, corresponding to relative risk from red signal of .5 to 1710. We modelled the number of participants by units of 10 from 20 to 100. For each combination of relative risk and number of participants, we created 1000 simulated datasets. Each simulated trial was analysed using logistic regression with a random intercept at the level of (simulated) study participant. The regression was in the form of equation 1:

$$dropout \sim \alpha + \beta_{red} * red + (1 \mid id)$$

Using a critical level of .05 for statistical significance, the power was estimated as the observed proportion of runs resulting in statistically significant fitted coefficients for the additional risk associated with the red signal.

The estimated power as a function of sample size and expected proportion of dropouts following the red signal are found in Figure 1. As the figure illustrates, power varies greatly as a function of sensitivity/relative risk. With 50 participants, we estimated 80% power to detect a true positive rate from the red signal of approximately 30%. As mentioned previously, the reported hit rate (sensitivity) of OQ from the red signal of bad treatment outcome is around 85%, leading us to expect a high relative risk of dropout following the red signal. Based on these estimates, we aimed for a sample size of approximately 50 respondents.

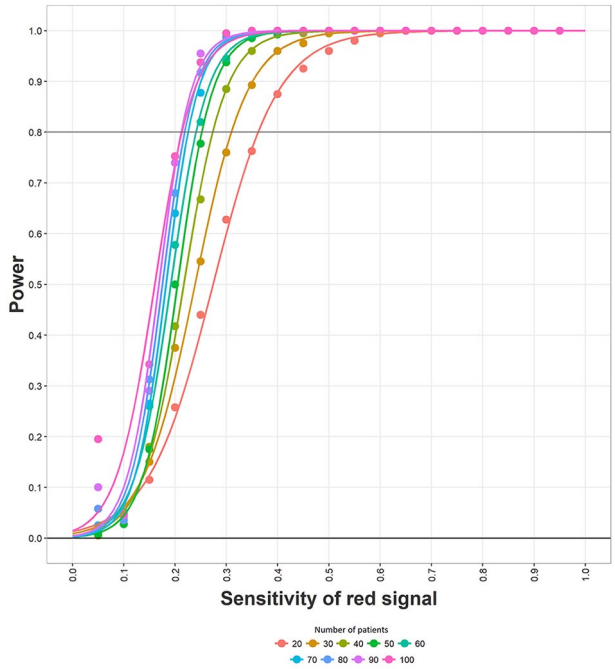

Figure 1. Simulation-based power estimates (y-axis) from runs with varying numbers of participants (delineated by colour), varying assumed sensitivity of 'red signal' to subsequent dropout (x-axis). Lines represent generalised linear models with a logit link, predicting observed powers for each number of simulated respondents by varying sensitivity.

Analyses of prediction. To estimate the predictive accuracy of the red signal for subsequent dropout, we used a mixed-effects logistic regression model with random intercepts at the level of individual patients. The model was specified to predict a dummy variable indicating dropout in the next seven days following an administered OQ questionnaire, with the variable of interest a separate dummy indicating the presence or absence of the red signal.

Our analysis included all dropout events, meaning that patients who returned to treatment within 90 days of leaving AMA could drop out more than once. This is in line with the intended use of the OQ-Analyst; the inclusion of a random intercept at the level of individual participants should account at least in part for individual differences in tendency to drop out of treatment. To ensure that the results were not overly influenced by many dropouts from a few individuals, we performed sensitivity analyses using the same methods, limiting the number of dropouts to the first per patient. Since using the first event only can be construed as a survival analysis, we also applied Cox regression with the red signal as a time-dependent covariate as a separate sensitivity analysis.

The OQ algorithm explicitly takes into account the longitudinal development of scores from each patient. Correspondingly, the clinical recommendations associated with the colour code outcomes from the OQ are independent of previous reports for the same patient; a red signal should be considered just as worrying for a patient previously reported to be on track as for a patient that has been off track for a long time. Conversely, if a patient previously indicated as at risk by the red signal is found to be on track, the clinical indication of risk should be the same as for a patient that has always been on track. Arguably, therefore, the most appropriate method of analysis would be to handle all observations as fully independent. The rationale may seem counterintuitive, but may be easier to grasp by use of an analogy: consider a case in which we use an instrument calibrated to indicate risk of cerebral stroke, accounting for various risk factors, including age. While we know that risk of stroke is highly age-dependent, this is already handled by the measure, such that a high indicated risk of stroke implies the same absolute risk, regardless of age. In such a case, further adjusting for age makes little sense, and considerations of predictive accuracy should be performed without including age as a separate predictor (unless the objective is to test whether age dependency is appropriately handled). In the case of OQ, previous response history is explicitly taken into account in the risk estimation algorithm. Consequently, testing the predictive accuracy of the instrument as it is intended to be used should theoretically consider each OQ risk indication in isolation. However, since the clinical populations on which the OQ algorithm has been calibrated may differ from substance abuse patients to some unknown extent, we take a more conservative approach and use; a mixed-model design with random intercepts at the level of individual patients to account for the lack of independence between responses made by the same individual.

The main analyses matches the analyses used in power calculation: a logistic regression model predicting observed dropout in the week following an administered OQ questionnaire, using a dummy variable indicating the red signal as the predictor of interest. To account for lack of independence between responses made by the same individual, we included a random intercept at the level of individual study participants.

As a measure of strength of evidence, we calculated Bayes factor between the null hypothesis (no difference in risk of dropout following red signal) and a range of assumed prior sensitivities of the red signal.

All analyses were performed in R 3.2.3 using Rstudio.[39,40] Mixed-effects models were run using the packages *lme4* (linear mixed-effects models)[41] and *brms* (Bayesian Multilevel Models using Stan).[42] Bayes factors were calculated using bespoke code, supported by the *brms* and BayesFactor packages.[43]

Since the algorithm used by the OQ-Analyst tracks changes from initial distress score, we performed sensitivity analyses in which initial distress score was included as a predictor and as a moderator for the red signal.

Results

Participants

For the 40 patients with the minimum required number of OQ responses, we had total of 647 OQ observations, see Figure 2 and Table 1.

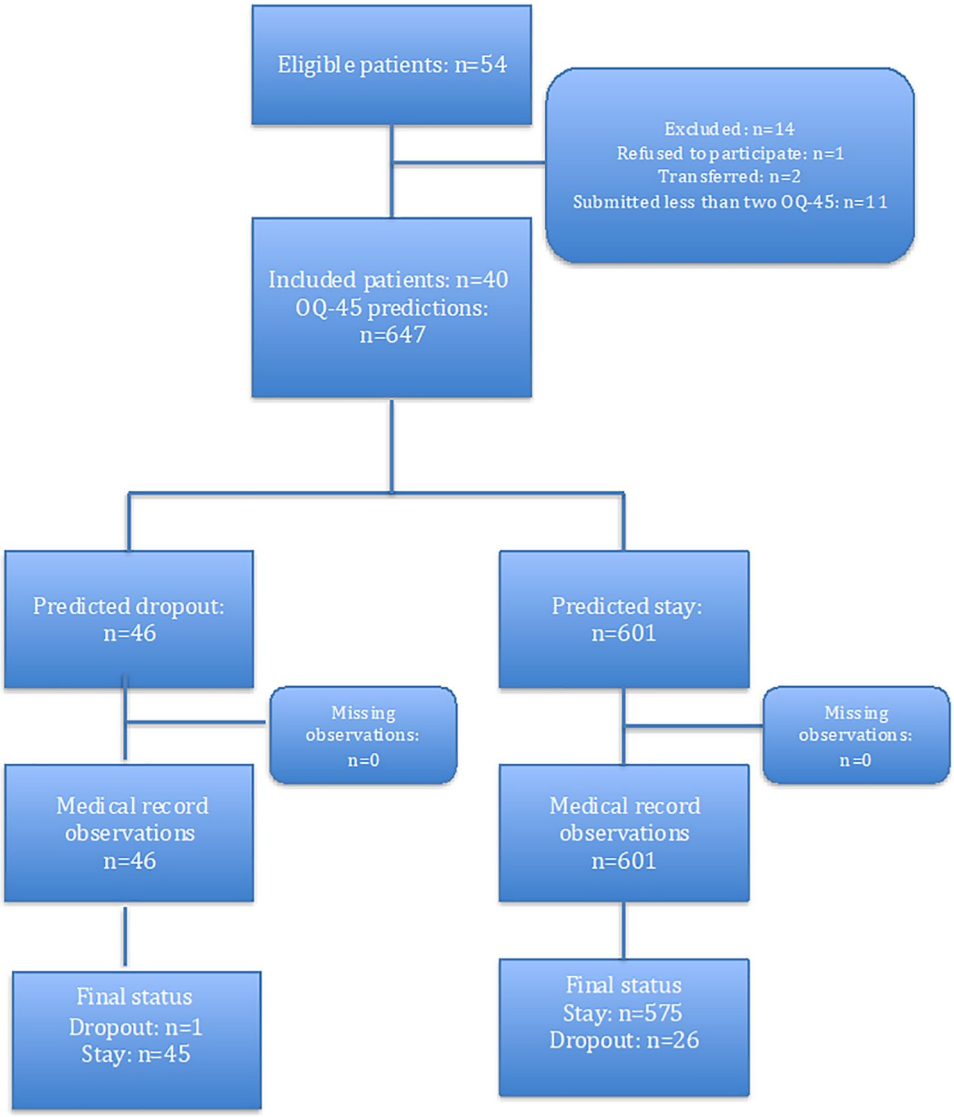

Figure 2. Flow chart compliant with STARD showing patient recruitment, OQ-Analyst predictions and observations extracted from medical journals.

Figure 3 displays the length of stay for the included patients, along with dropouts, red signal and indications of completed treatment.

There were a total of 46 red signals, and 32 dropouts of at least 1 week, 27 of which followed after a valid administration of OQ, that is after completing the minimum requirement of two OQs. Of the 27, only one followed a red signal. Figure 3 displays red signals and subsequent dropout. Results from the fitted mixed-effects logistic regression model can be found in Table 2, which indicates a statistically nonsignificant, lower risk of dropout following a red signal than other OQ responses.

In the sensitivity analysis using Cox regression with a red signal as a time-dependent covariate, the red signal was found to be associated with a statistically significant ($P < .001$) reduction in proportional risk of subsequent dropout.

Figure 4 displays the raw factors (BR01) for different assumed priors regarding true positives for dropouts following a red signal, and indicates that no association between red signal and subsequent dropout is substantially more likely than priors for relative risks above 2 for a red signal (true positives rates over 13.5%).

We performed a large number of sensitivity analyses using 0 day (any unplanned absence), 1 week (⩾7 days), 2 weeks (⩾14 days), and 1 month (⩾35 days) as cut-offs. Anonymised research data are available as supplemental material (S3). Sensitivity analyses with shorter or longer required absence to be counted as a dropout altered the number of observed dropouts following valid OQ responses (54 dropouts of any duration to 16 dropouts of at least 90 days). Point estimates for the risk associated with red signal were consistently negative for all definitions.

The sensitivity analyses including baseline score as a predictor and as a moderator for red signal did not alter the findings provided by our main analyses.

Discussion

The general aim of the present study was to investigate the ability of the OQ-Analyst to predict dropout from an inpatient

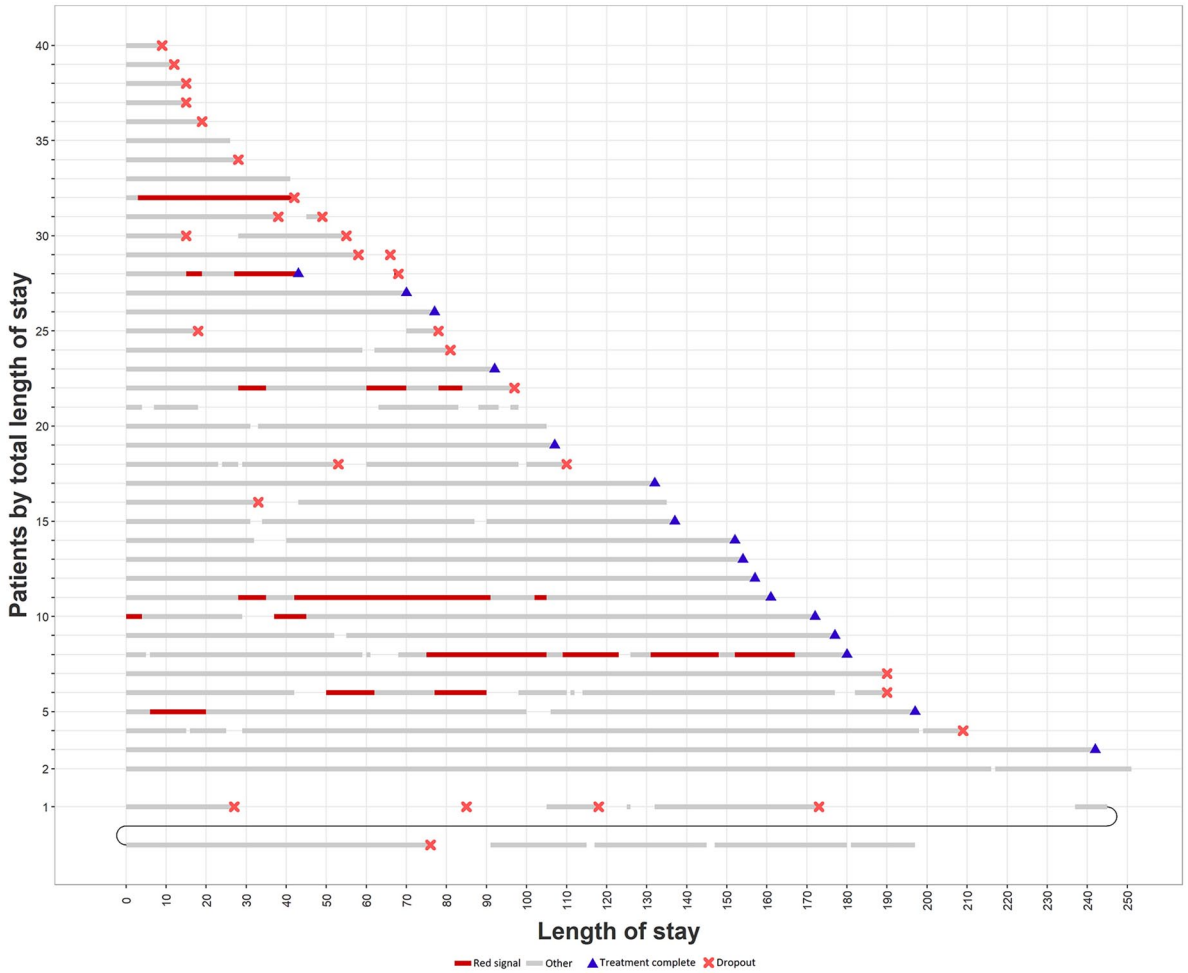

Figure 3. Length of stay (x-axis) for included patients, sorted by total length of stay. Red x denotes dropout, blue triangle completed treatment. Red reflects red signal, grey otherwise. No line indicates that patient was not admitted, for example, due to dropout or out on leave. Treatment cessation with no marker due to any other reason, such as the patient being moved to another ward.

Table 2. Random intercept logistic regression predicting dropout from red signal.

	ESTIMATE	CI [2.5%, 97.5%]	SE	*P*
A. ALL DROPOUTS INCLUDED				
Intercept	−3.587	−4.746, −2.838	.448	.000
Red signal	−.660	−3.814, 1.420	1.227	.591
B. FIRST DROPOUT PER PARTICIPANT				
Intercept	−3.537	−4.019, −3.055	.246	.000
Red signal	−.270	−2.310, 1.770	1.041	.793

SUD treatment. Dropout is often linked to deterioration,[44] and the OQ-Analyst use the same prediction model to predict both outcomes. Based on studies addressing the OQ-Analyst ability to predict deterioration,[29–34] we hypothesised that the system would be able to predict a more frequently occurring event such as dropout, above the level of deterioration, that is above a hit rate of 85%. The results indicated that the OQ-Analyst was unable to predict dropout on a week-by-week basis and as a final outcome.

We think there are at least four reason that might explain the lack of prediction in our study. First, our study was not powered to recognise an increased relative risk of dropout less than 2, meaning that the OQ-Analyst may be able to predict dropout below this level. However, the practical value of prediction at this level is questionable. Consider an example: an average sized inpatient unit in Norway treats about 50 patients per year, with approximately 25 dropouts. Assuming that the length of stay (approximately 16 weeks) and proportion of OQ

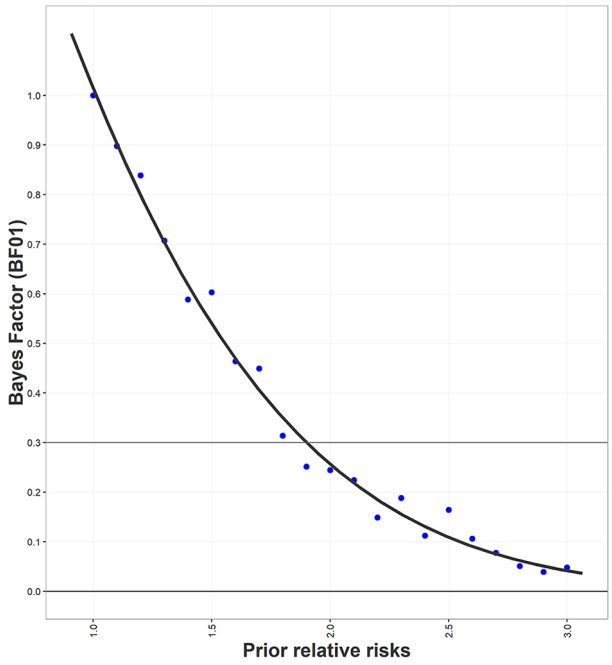

Figure 4. Simulation-based Bayes factor estimates based on assumed prior relative risk of dropout following 'red signal' and the observed patient data. The line is a fitted exponential curve on risk. BF01 falls below with assumed prior relative risks exceeding approximately 1.9.

responses yielding a red signal (approximately 7%) observed in this study is representative, a relative risk of 2 would mean that the OQ-Analyst, in the best case, correctly predicts two dropouts per year for the average clinic. As the instrument is intended for individual decision making, the system needs to detect differences large enough to make a practical, noticeable effect in clinical daily life. Additionally, the OQ produced 46 red signals not followed by dropout in this sample, misidentifying the patients as likely to drop out. In a clinical setting, this might result in unnecessary interventions and expenditure of additional resources at the cost of other tasks and patients. There is call for a debate on what levels of sensitivity and specificity should be required of a clinical prediction tool before it can be recommended for use in clinical practice.

A second explanation for the finding could be that our sample was biased toward the higher end of disturbance. Poorer functioning patients are more likely to show a mismatch with general outcome measures such as the OQ-Analyst.[23,45,46] However, when we performed the analyses with initial distress score as moderator we did not detect any systematic patterns of differentiated risk. With regard to other potentially biasing factors such as gender,[47] prevalence of co-occurring mental disorders,[48] patterns of drug use[49] and dropout rate,[8] our sample seemed reflective of international SUD populations, thus making our sample representative for other SUD populations.

A third possible explanation for the lack of prediction may be due to different mechanisms underlying dropout from mental health services and dropout from SUD treatment. The OQ-Analyst was originally developed for mental health, and based on the test items included in the OQ-Analyst, dropout can be predicted from items tracking three broad areas of adult mental health: symptom distress; interpersonal problems; and functioning in everyday life. Conclusions from a recent study exploring potential mechanisms of dropout from SUD treatment[50] correspond well with the domains covered by the OQ-Analyst. However, the study highlighted the importance of drug craving, a symptom currently not found in the OQ-Analyst.

Lastly, the lack of prediction could be the result of a mismatch between the data on which the OQ-Analyst is constructed on and SUD populations. The OQ-Analyst uses intake score and deviation from expected treatment response to predict dropout. Both are reported to differ between mental health and SUD populations, with SUD patients tending to report significantly less distress at intake, and also needing longer time in treatment before progress occur.[51]

The lack of prediction indicates that the OQ-Analyst may not be appropriate for SUD patients and underlines the need for validation studies of PM systems used in mental health for SUD populations. Where one might be concerned that the findings were caused by a few individuals dropping out repeatedly, the sensitivity analyses using the first dropout per patient only were no more favourable to a red signal as an indicator of subsequent dropout. To the contrary, the Cox regression indicated a significantly reduced proportional hazard rate for individuals following a red signal. In the absence of empirical evidence, we cannot recommend the OQ-Analyst for use with the purpose of predicting dropout from inpatient SUD treatment. It should be noted that it is possible that the instrument provides the therapist with other meaningful information leading to improved retention and research on the preventive ability of the OQ-Analyst in a SUD treatment setting is needed.

The failure to predict dropout in the present study is uninformative as to the ability of the OQ-Analyst to predict dropout in other patient groups. The concept behind the claim of having such ability is intuitively appealing: the OQ-Analyst predicts poor progress;[29–34] poor progress is a core element in dropout, thus a method used to predict poor progress should predict dropout. However, dropout and poor progress may be different types of negative outcome. This has been reported in previous research,[52] and Linden (2013)[53] suggests that deterioration is an adverse reaction to the treatment, while dropout may or may not be treatment related. This understanding indicates that prediction failure implies model inadequacy, which could mean that the OQ-Analyst is unable to predict dropout regardless of patient population. There is call for studies testing the ability of the OQ-Analyst to predict dropout in other patient groups.

Strengths and limitations

This study has several strengths. First, it was performed in a naturalistic setting with less restrictive methodological standards in terms of patient selection, therapist competence and

adherence, and other issues relating to design, thus making our result better generalisable to real-world settings. Moreover, this study used electronic medical records as the standard for predictive accuracy, whereas previous studies have used the OQ-Analyst as a reference standard for predictions made by the OQ-Analyst. External validation is necessary as prediction models tend to perform better on data on which the model is constructed on compared to new data.[54]

Several limitations of the study should also be noted. While sufficient to detect predictive ability of magnitudes such as reported for negative treatment outcomes, the sample size in this study implies insufficient power to detect relative risks less than 2 for dropouts following a red signal. However, the null finding above this level is trustworthy as demonstrated by the calculated Bayes factor.

Another potential source of error could result from the study design. In order to investigate predictive accuracy, we had to withhold the feedback from the therapists, thus withholding the potential benefits of undergoing repeated testing for our participating patients. As the therapists were unable to detect bogus responses, this could have made our data vulnerable to erroneous answers from unmotivated participants. We attempted to mend this issue by generating a reliability test for items in the OQ-Analyst that should be answered differently (e.g. I get along well with others should be answered differently from I often get in conflict with others). The test flagged assessments with inconsistent answers, which were then disregarded in the analyses.

Conclusion

There are two potential main implications in this study. First, the results suggest that the dropout predictor in the OQ-Analyst does not predict subsequent dropout for SUD patients in inpatient treatment and indicates a need for disorder-specific PM instruments rather than general instruments such as the OQ-Analyst. Second, the OQ-Analyst's ability to predict dropout appears not to be documented in general. In the absence of empirical evidence, it may be safer not to assume such ability.

Overall, the result emphasises the importance of, and need for, studies designed to directly test the OQ-Analysts dropout predictor. Until empirical evidence is available to suggest the ability of OQ-Analyst to predict dropout, statements about having such qualities should be made with care. We are happy to notice that the company distributing the OQ-Analyst has, sometime in the last few months, removed all claims of predicting dropout from their web page, along with removing dropout from the feedback message following a red signal.

Acknowledgements

We thank Erik Vasaasen and Espen Walderhaug for technical support.

ORCID iD

Hanne H Brorson https://orcid.org/0000-0001-6500-4035

Supplemental Material

Supplemental material for this article is available online.

REFERENCES

1. Alterman AI, McKay JR, Mulvaney FD, McLellan AT. Prediction of attrition from day hospital treatment in lower socioeconomic cocaine-dependent men. *Drug Alcohol Depend*. 1996;40(3):227–233.
2. Brewer DD, Catalano RF, Haggerty K, Gainey RR, Fleming CB. RESEARCH REPORT A meta-analysis of predictors of continued drug use during and after treatment for opiate addiction. *Addiction*. 1998;93(1):73–92.
3. Moos RH, Pettit B, Gruber VA. Characteristics and outcomes of three models of community residential care for abuse patients. *J Subst Abuse*. 1995;7(1):99–116.
4. Simpson DD. Treatment for drug abuse: Follow-up outcomes and length of time spent. *Arch Gen Psychiatry*. 1981;38(8):875–880.
5. United Nations Office on Drugs and Crime. *World Drug Report 2012*. Austria: United Nations; 2012. http://www.unodc.org/unodc/en/data-and-analysis/WDR-2012.html. Accessed November 30, 2012.
6. Dalsbø TK, Hammerstrøm KT, Vist GE, et al. Psychosocial interventions for retention in drug abuse treatment. *Cochrane Database Syst Rev*. 2010;(1):CD008220. doi:10.1002/14651858.CD008220
7. Ravndal E, Vaglum P, Lauritzen G. Completion of long-term inpatient treatment of drug abusers: a prospective study from 13 different units. *European addiction research*. 2005;11(4):180.
8. Brorson HH, Ajo Arnevik E, Rand-Hendriksen K, Duckert F. Drop-out from addiction treatment: a systematic review of risk factors. *Clin Psychol Rev*. 2013;33(8):1010–1024.
9. Stark MJ. Dropping out of substance abuse treatment: a clinically oriented review. *Clin Psychol Rev*. 1992;12(1):93–116.
10. Craig RJ. Reducing the treatment drop out rate in drug abuse programs. *J Subst Abuse Treat*. 1985;2(4):209–219.
11. Baekeland F, Lundwall L. Dropping out of treatment: a critical review. *Psychol Bull*. 1975;82(5):738–783.
12. Cacciola JS, Alterman AI, Dephilippis D, et al. Development and initial evaluation of the Brief Addiction Monitor (BAM). *J Subst Abuse Treat*. 2013;44(3):256–263.
13. Lambert MJ. Helping clinicians to use and learn from research-based systems: the OQ-analyst. *Psychotherapy*. 2012;49(2):109.
14. Pulford J, EDeering D, Robinson G, et al. Development of a routine outcome monitoring instrument for use with clients in the New Zealand alcohol and other drug treatment. Sector: the Alcohol and Drug Outcome Measure (ADOM). (The Drug and Alcohol Outcome Measure (ADOM)). 2010. http://www98.griffith.edu.au/dspace/handle/10072/36053. Accessed August 30, 2014.
15. Howard KI, Kopta Mark S, Krause MS, Orlinsky DE. The dose–effect relationship in psychotherapy. *Am Psychol*. 1986;41(2):159–164.
16. McCaul ME, Svikis DS, Moore RD. Predictors of outpatient treatment retention: patient versus substance use characteristics. *Drug Alcohol Depend*. 2001;62(1):9–17.
17. Ackerman SJ, Benjamin LS, Beutler LE, et al. Empirically supported therapy relationships: conclusions and recommendations fo the Division 29 Task Force. *Psychother Theor Res Pract Train*. 2001;38(4):495–497.
18. Lambert MJ, Hansen NB, Finch AE. Patient-focused research: using patient outcome data to enhance treatment effects. *J Consult Clin Psychol*. 2001;69(2):159.
19. Whipple JL, Lambert MJ, Vermeersch DA, Smart DW, Nielsen SL, Hawkins EJ. Improving the effects of psychotherapy: the use of early identification of treatment failure and problem-solving strategies in routine practice. *J Couns Psychol*. 2003;50(1):59–68.
20. Duncan BL, Miller SD, Sparks JA. *The Heroic Client: A Revolutionary Way to Improve Effectiveness through Client-Directed, Outcome-Informed Therapy*. Hoboken, NJ: John Wiley & Sons; 2004.
21. Miller SD, Duncan BL, Brown J, Sorrell R, Chalk MB. Using formal client feedback to improve retention and outcome: making ongoing, real-time assessment feasible. *Journal of Brief Therapy*. 2006;5(1):5–22.
22. Goodman JD, McKay JR, DePhilippis D. Progress monitoring in mental health and addiction treatment: a means of improving care. *Prof Psychol Res Pract*. 2013;44(4):231–246.
23. Conijn JM, Emons WHM, De Jong K, Sijtsma K. Detecting and explaining aberrant responding to the outcome questionnaire-45. *Assessment*. 2015;22(4):513–524.
24. Lambert MJ. *Administration and Scoring Manual for the OQ-45.2 (Outcome Questionnaire)*. OQ Measures, LLC; 2004.
25. Hatfield DR, Ogles BM. The use of outcome measures by psychologists in clinical practice. *Prof Psychol Res Pract*. 2004;35(5):485.
26. Samantha.Elliott. NREPP. https://www.samhsa.gov/nrepp. Published May 27, 2014. Accessed June 14, 2018.

27. Lambert MJ, Whipple JL, Vermeersch DA, et al. Enhancing psychotherapy outcomes via providing feedback on client progress: a replication. *Clin Psychol Psychother*. 2002;9(2):91–103.
28. Lambert M. Presidential address: what we have learned from a decade of research aimed at improving psychotherapy outcome in routine care. *Psychother Res*. 2007;17(1):1–14.
29. Hannan C, Lambert MJ, Harmon C, et al. A lab test and algorithms for identifying clients at risk for treatment failure. *J Clin Psychol*. 2005;61(2):155–163.
30. Ellsworth JR, Lambert MJ, Johnson J. A comparison of the outcome questionnaire-45 and outcome questionnaire-30 in classification and prediction of treatment outcome. *Clin Psychol Psychother*. 2006;13(6):380–391.
31. Lambert MJ, Whipple JL, Bishop MJ, Vermeersch DA, Gray GV, Finch AE. Comparison of empirically-derived and rationally-derived methods for identifying patients at risk for treatment failure. *Clin Psychol Psychother*. 2002;9(3):149–164.
32. Lutz W, Lambert MJ, Harmon SC, Tschitsaz A, Schürch E, Stulz N. The probability of treatment success, failure and duration – what can be learned from empirical data to support decision making in clinical practice? *Clin Psychol Psychother*. 2006;13(4):223–232.
33. Percevic R, Lambert MJ, Kordy H. What is the predictive value of responses to psychotherapy for its future course? Empirical explorations and consequences for outcome monitoring. *Psychother Res*. 2006;16(3):364–373.
34. Spielmans GI, Masters KS, Lambert MJ. A comparison of rational versus empirical methods in the prediction of psychotherapy outcome. *Clin Psychol Psychother*. 2006;13(3):202–214.
35. Amble I, Gude T, Stubdal S, et al. Psychometric properties of the outcome questionnaire-45.2: the Norwegian version in an international context [published online ahead of print November 5, 2013.]. *Psychother Res*. doi:10.1080/10503307.2013.849016
36. Snell MN, Mallinckrodt B, Hill RD, Lambert MJ. Predicting counseling center clients' response to counseling: a 1-year follow-up. *J Couns Psychol*. 2001;48(4):463–473.
37. Vermeersch DA, Lambert MJ, Burlingame GM. Outcome questionnaire: item sensitivity to change. *J Pers Assess*. 2000;74(2):242–261.
38. Alfandre DJ. 'I'm Going Home': discharges against medical advice. *Mayo Clin Proc*. 2009;84(3):255–260.
39. R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2015. https://www.R-project.org/; https://www.google.com/search?q=R+Core+Team+%282015%29.+R%3A+A+language+and+environment+for+statistical+computing.+R+++Foundation+for+Statistical+Computing%2C+Vienna%2C+Austria.+URL+++https%3A%2F%2Fwww.R-project.org%2F.&ie=utf-8&oe=utf-8&client=firefox-b. Accessed September 18, 2018.
40. RStudio Team. RStudio: Integrated Development for R. Boston, MA: RStudio, Inc., 2015. http://www.rstudio.com/; https://www.google.com/search?q=RStudio+Team+%282015%29.+RStudio%3A+Integrated+Development+for+R.+RStudio%2C+Inc.%2C+Boston%2C+MA+URL+http%3A%2F%2Fwww.rstudio.com%2F.&ie=utf-8&oe=utf-8&client=firefox-b. Accessed September 18, 2018.
41. Bates D, Maechler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Softw*. 2015;67(1):1–48.
42. Bürkner P-C. brms: An R Package for Bayesian Multilevel Models Using Stan. *J Stat Softw*. 2017;80(1):1–28.
43. Morey RD, Rouder JN. BayesFactor: Computation of Bayes Factors for Common Designs. R package version 0.9.12-4.2. 2018. https://CRAN.R-project.org/package=BayesFactor. https://www.google.com/search?q=Richard+D.+Morey+and+Jeffrey+N.+Rouder+%282018%29.+BayesFactor%3A+Computation+of+Bayes+++Factors+for+Common+Designs.+R+package+version+0.9.12-4.2.+++https%3A%2F%2FCRAN.R-project.org%2Fpackage%3DBayesFactor&ie=utf-8&oe=utf-8&client=firefox-b. Accessed September 18, 2018.
44. Slone NC, Reese RJ, Mathews-Duvall S, Kodet J. Evaluating the efficacy of client feedback in group psychotherapy. *Group Dyn Theor Res Pract*. 2015;19(2):122.
45. Conijn JM, Emons WHM, van Assen MALM, Pedersen SS, Sijtsma K. Explanatory, Multilevel Person-Fit Analysis of Response Consistency on the Spielberger State-Trait Anxiety Inventory. *Multivariate Behav Res*. 2013;48(5):692–718.
46. Woods CM, Oltmanns TF, Turkheimer E. Detection of aberrant responding on a personality scale in a military sample: an application of evaluating person fit with two-level logistic regression. *Psychol Assess*. 2008;20(2):159–168.
47. Substance Abuse and Mental Health Services Administration. *Results from the 2013 National Survey on Drug Use and Health: Summary of National Findings*, NSDUH Series H-48, HHS Publication No. (SMA) 14-4863. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2014.
48. Grant BF, Stinson FS, Dawson DA, et al. Prevalence and co-occurrence of substance use disorders and independent mood and anxiety disorders: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Arch Gen Psychiatry*. 2004;61(8):807–816.
49. Kedia S, Sell MA, Relyea G. Mono- versus polydrug abuse patterns among publicly funded clients. *Subst Abuse Treat Prev Policy*. 2007;2:33.
50. Nordheim K, Walderhaug E, Alstadius S, Kern-Godal A, Arnevik E, Duckert F. Young adults' reasons for dropout from residential substance use disorder treatment. *Qual Soc Work*. 2018;17(1):24–40.
51. Miller SD, Duncan BL, Sorrell R, Brown GS. The partners for change outcome management system. *J Clin Psychol*. 2005;61(2):199–208.
52. Saxon D, Barkham M, Foster A, Parry G. The Contribution of Therapist Effects to Patient Dropout and Deterioration in the Psychological Therapies. *Clin Psychol Psychother*. 24(3):575–588.
53. Linden M. How to define, find and classify side effects in psychotherapy: from unwanted events to adverse treatment reactions. *Clin Psychol Psychother*. 20(4):286–296.
54. Bleeker SE, Moll HA, Steyerberg EW, et al. External validation is necessary in prediction research: a clinical example. *J Clin Epidemiol*. 2003;56(9):826–832.