Received: 15 May 2019 | Revised: 2 January 2020 | Accepted: 3 March 2020

DOI: 10.1002/cpp.2441

RESEARCH ARTICLE

WILEY

Why are some cases not on track? An item analysis of the Assessment for Signal Cases during inpatient psychotherapy

Thomas Probst[1] | Maria Kleinstäuber[2] | Michael J. Lambert[3] | Karin Tritt[4] | Christoph Pieh[1] | Thomas H. Loew[4] | Reiner W. Dahlbender[5] | Jaime Delgadillo[6]

[1]Department for Psychotherapy and Biopsychosocial Health, Danube University Krems, Krems an der Donau, Austria

[2]Department of Psychological Medicine, Dunedin School of Medicine, University of Otago, Dunedin, New Zealand

[3]Department of Psychology, Brigham Young University, Provo, Utah

[4]Department of Psychosomatic Medicine, University Hospital Regensburg, Regensburg, Germany

[5]Clinic for Psychosomatic Medicine and Psychotherapy, University Hospital Ulm, Ulm, Germany

[6]Clinical Psychology Unit, Department of Psychology, University of Sheffield, Sheffield, UK

Correspondence
Thomas Probst, Department für Psychotherapie und Biopsychosoziale Gesundheit Donau-Universität Krems, Dr.-Karl-Dorrek-Straße 30, Krems an der Donau 3500, Austria.
Email: thomas.probst@donau-uni.ac.at

Abstract

Within the Routine Outcome Monitoring system "OQ-Analyst," the questionnaire "Assessment for Signal Cases" (ASC) supports therapists in detecting potential reasons for not-on-track trajectories. Factor analysis and a machine learning algorithm (LASSO with 10-fold cross-validation) were applied, and potential predictors of not-on-track classifications were tested using logistic multilevel modeling methods. The factor analysis revealed a shortened (30 items) version of the ASC with good internal consistency (α = 0.72–0.89) and excellent predictive value (area under the curve = 0.98; positive predictive value = 0.95; negative predictive value = 0.94). Item-level analyses showed that interpersonal problems captured by specific ASC items (not feeling able to speak about problems with family members; feeling rejected or betrayed) are the most important predictors of not-on-track trajectories. It should be considered that our results are based on analyses of ASC items only. Our findings need to be replicated in future studies including other potential predictors of not-on-track trajectories (e.g., changes in medication, specific therapeutic techniques, or treatment adherence), which were not measured this study.

KEYWORDS

Assessment for Signal Cases, progress feedback, psychotherapy, routine outcome monitoring

1 | INTRODUCTION

Research has repeatedly shown that not all mental health patients benefit from psychotherapy. Approximately in 5–10% of adult patients (Lambert, 2013a) and in 14–24% of child and adolescent patients (Warren, Nelson, Burlingame, & Mondragon, 2012), mental health deteriorates during the course of psychological treatment. A number of Routine Outcome Monitoring (ROM) systems (Boswell, Kraus, Miller, & Lambert, 2015) have been developed to support patient-focused psychotherapy research (Howard, Moras, Brill, Martinovich, & Lutz, 1996). These systems aim to systematically monitor relevant indicators of patients' mental health (e.g., psychological symptoms) and to provide feedback about patients' progress to clinicians (and patients) throughout the course of therapy. ROM can help to identify patients with negative symptom trajectories early in the course of therapy and to counteract treatment failure. One of the most widely used ROM system is the OQ-Analyst (Lambert, 2012, 2015), which helps to track patients' weekly progress

The work was conducted at the University Hospital Regensburg (Psychosomatics Department of the Hospital in Donaustauf) and the Hospital "Am schönen Moos" Bad Saulgau, Germany.

using the Outcome Questionnaire-45 (OQ-45, Lambert et al., 2004; Haug, Puschner, Lambert, & Kordy, 2004; Lambert, Hannöver, Nisslmüller, Richard, & Kordy, 2002). The OQ-45 measures three domains including symptoms (e.g., anxiety and depression), interpersonal problems, and social role functioning. The total score is a measure of patients' general mental health functioning. The OQ-Analyst software plots expected recovery curves for patient groups with differing levels of intake OQ-45 scores. This allows the identification of cases with extreme deviations in the OQ-45 by comparison with an expected trajectory of improvement, starting with the second therapy session. These expected treatment response curves were based on a sample of over 11,000 patients who completed the OQ-45 during the course of treatment in a variety of routine care clinical settings (Finch, Lambert, & Schaalje, 2001). According to Lambert (2013b), 40% of the outcome variance at the end of therapy is explained by the OQ-45 intake score and changes in OQ-45 scores from Sessions 1 to 3. Further variables such as demographic or diagnostic information can explain incremental variance of only 1%. Accordingly, the OQ-Analyst predictions of expected treatment response are simply based on the initial OQ-45 score at the start of therapy.

Previous research has shown that patients whose OQ-45 score deviates negatively from their expected trajectory at least once during treatment ("not-on-track" [NOT] signal) are at risk of treatment failure (Hannan et al., 2005; Lambert, Whipple et al., 2002; Spielmans, Masters, & Lambert, 2006). Lutz et al. (2006) reported that the probability of treatment failure increases as a function of the number of NOT signals in patients' trajectory of change. Clinical support tools (CSTs) are implemented in the OQ-Analyst to support therapists in improving outcomes for patients with NOT signals during treatment (Harmon, Hawkins, Lambert, Slade, & Whipple, 2005; Lambert et al., 2007). These CSTs comprise a questionnaire called the Assessment for Signal Cases (ASC; Lambert et al., 2007) with a linked clinical decision tree, as well as intervention handouts to support clinicians in problem solving. The ASC consists of 40 items (answered on a 5-point Likert scale), which aim to assess the following four domains: therapeutic alliance, social support, motivation, and negative life events. These domains were chosen with reference to psychotherapy process variables highlighted as important in previous studies (therapeutic alliance: e.g.,Flückiger, Del Re, Wampold, & Horvath 2018; social support: e.g., Roehrle and Strouse (2008); motivation: e.g., Norcross, Krebs, and Prochaska 2011; life events: e.g., Sexton 1996). In a review of psychotherapy research, Asay and Lambert (1999) estimated that factors that are external to the therapy process (e.g., social support and critical life events) may explain up to 40% of therapy outcome, therapeutic alliance up to 30%, patients' therapy motivation and expectations up to 15%, and the remaining 15% may be explained by other psychotherapeutic models and interventions, which are not captured by the ASC.

In clinical practice, the ASC is intended to be administered only when patients' symptom trajectory is NOT to support clinicians in assessing if poor progress may be related to problems in one or more of the four ASC domains. In one controlled trial, the ASC was provided to all patients (NOT and on track) on a weekly basis in order to monitor their ASC domains continuously throughout the course of therapy (Probst et al., 2013; Probst, Lambert, Dahlbender, Loew, & Tritt, 2014; Probst, Lambert, Loew, Dahlbender, & Tritt, 2015). This trial replicated the results of previous studies: Progress feedback based on the OQ-45 combined with CSTs improves the outcome of NOT cases (see also Harmon et al., 2007; Lambert, Whipple, & Kleinstauber, 2018; Shimokawa, Lambert, & Smart, 2010; Slade, Lambert, Harmon, Smart, & Bailey, 2008; Whipple et al., 2003). Probst et al. (2015) found that the social support and negative life events domains of the ASC were associated with extremely negative deviations from expected treatment response trajectories (NOT signals). The important role of social support in NOT cases, compared with other ASC domains, was also demonstrated by White et al. (2015). These findings suggest that the therapeutic alliance and motivation domains of the ASC are rather unrelated to NOT signals. An alternative explanation may be that the ASC domains might include single items that have weak correlations with other items, thus undermining the factorial validity of the measure and introducing "noise" in outcome prediction analyses. It is possible that highly specific aspects of the ASC domains (e.g., features captured by single items) could be related to NOT signals, but such correlations could be obscured by entering other "noise" items in analyses. Moreover, the four-factor structure of the ASC has been proposed by the authors of this measure but has not been tested with psychometric methods yet, which casts some doubt over the validity of findings of previous studies.

The present study aimed to extend our understanding of the aspects associated with poor progress in psychotherapy. The central research question of this study was "which items and domains of the ASC predict NOT signals in an inpatient clinical sample?" To answer this question, we examined the predictive value of the ASC domains proposed by authors of the measure compared with ASC domains extracted with factor analysis. Furthermore, we examined the predictive value of specific ASC items.

Key Practitioner Message

- Interpersonal problems were identified as important predictors of patients' poor response to therapy.
- The proposed four-factor structure of the Assessment for Signal Cases (ASC) was empirically supported.
- The predictive value of a shortened 30-item version of the ASC was as high as that of the 40-item ASC.
- Using the 30-item ASC is recommended for future research, given its brevity and predictive value.
- Special attention should be paid to two specific ASC items capturing interpersonal problems.

2 | METHODS

2.1 | Design, setting, and interventions

This study was based on the reanalysis of data from a randomized controlled trial on the efficacy of the OQ-Analyst in inpatient treatments for patients with severe psychosomatic problems (Probst et al., 2013, 2014). As typical for psychosomatic medicine in Germany (see Zipfel, Herzog, Kruse, & Henningsen, 2016), multiprofessional teams (mostly psychologists, physicians, and nurses) provided multimodal and multimethod treatments based on the biopsychosocial model comprising individual and group psychotherapy, relaxation or mindfulness training, physical activity therapy, art, dance, and music therapy and—if required—crisis intervention, visitations by nurses, and medical consultations. Inpatient psychotherapy often is recommended when outpatient treatment is deemed insufficient. Compared with patients that are referred to outpatient psychotherapy, patients referred to inpatient psychotherapy, for example, have longer sick leaves, are less able to work, exhibit more often a somatoform or personality disorder, are more burdened by psychological symptoms, have a lower functional level, and the personality structure is less favorable (Huber, Brandl, Henrich, & Klug, 2002). The core competencies of German psychosomatic clinics lie in treating somatoform or functional disorders, eating disorders, somatopsychic disorders (including psycho-oncology, psychocardiology, neuropsychosomatics, and psychodiabetology), as well as psychotraumatology and an overlap with psychiatry exists in depressive, anxiety, and personality disorders (Zipfel et al., 2016, p. 262). The patients participating in the trial that were reanalyzed here were randomly assigned to either inpatient treatment-as-usual or inpatient treatment-as-usual combined with feedback-informed treatment. Both groups were monitored on a weekly basis using standardized questionnaires described below. For those randomized to the feedback group, the responsible individual therapist received weekly feedback reports of the OQ-Analyst (Lambert, 2012). Further details about the trial such as the flowcharts are described by Probst et al. (2013, 2014). The study was approved by the Ethics Committee of the University Clinic of Regensburg, Germany.

2.2 | Measures

The OQ-45 (Lambert et al., 2004) is a self-report measure of psychological distress covering three domains of symptoms, interpersonal problems, and social functioning. The German version of the OQ has been reported to have good internal consistency (α = 0.93) and adequate retest reliability (r = 0.88; Lambert, Hannöver, et al., 2002). Each item is scored using a 5-point Likert scale, yielding a total severity score ranging from 0 to 180. The OQ-Analyst software (Lambert, 2012) uses well-established statistical methods that provide an automated NOT signal when a patient's responses to the OQ-45 indicate atypically high levels of distress, by comparison with cases with similar intake severity scores (Finch et al., 2001). This risk signal is intended to prompt the therapist to identify and to attempt to resolve problems that impede treatment progress.

The ASC (Lambert et al., 2007) is a 40-item self-report questionnaire covering four domains: therapeutic alliance (11 items, range of total score: 11 to 55), social support (11 items, range of total score: 11 to 55), motivation (nine items, range of total score: 9 to 45), and negative life events (nine items, range of total score: 9 to 45). These four areas have been proposed to influence patients' progress in psychotherapy and form the basis for clinical recommendations and techniques that may help to resolve obstacles impeding patients' treatment progress (Lambert et al., 2007). The ASC used in this study was translated into German via a back-translation method, yielding adequate reliability indices for each subdomain (α = 0.71 to 0.89; Probst et al., 2013). Some ASC items are reverse scored so that low domain scores indicate problems in each domain.

2.3 | Participants characteristics

Patient progress can only be classified as on track or NOT for treatment weeks following intake (NOT signal: Yes or No). At least one ASC assessment in these weeks is required to study associations between the ASC and NOT signals. Relevant data were available for 283 patients (receiving either treatment-as-usual or treatment-as-usual + feedback-informed treatment; data from both conditions were pooled to have a larger sample size). The majority of these patients was female (59.7%). On average, the participants were 48.27 (*SD* = 13.44) years old, and their intake OQ-45 score amounted to 81.11 (*SD* = 25.67). The most frequent psychiatric diagnoses according to ICD-10 were depressive disorders (F32: 10.3%; F33: 19.6%), somatoform disorders (F45: 29.1%), anxiety-related disorders (F40: 2.8%; F41: 9.1%; F42: 1.6%; F43: 7.2%), and eating disorders (F50: 2.3%; F51: 3.1%). The average amount of psychiatric diagnoses per patient was 2.25 (*SD* = 1.09).

2.4 | Statistical analysis

We approached our overall aim of the study—to investigate the information of the ASC that is most important to predict NOT signals—using factor analysis and signal detection methods. The statistical analysis was organized in three stages guided by specific objectives: (a) to examine the factor structure and reliability of the ASC measure domains; (b) to identify specific ASC items that are associated with NOT signals; (c) to compare the predictive value of factor analysis versus signal detection approaches.

Stage 1 applied factor analysis using all 40 ASC items from the intake assessment. Missing data for cases (*N* = 32; 11.3%) that did not respond to an ASC item were imputed by averaging the values from 25 estimated datasets using an expectation maximization method (Schafer & Olsen, 1998). We initially examined the adequacy of the dataset for factor analysis using the Kaiser–Meyer–

Olkin (KMO) test and Bartlett's test of sphericity. In order to empirically determine how many factors optimally explained the variability in the data, we applied parallel analysis using polychoric correlations, based on unweighted least squares, with promin rotation (Lorenzo-Seva, 1999; Timmerman & Lorenzo-Seva, 2011). Items that had factor loadings <.50 were excluded, in order to retain a parsimonious set of items that were strongly associated with each domain. Cronbach's alpha was used to examine the internal consistency of each domain.

Taking a different approach in Stage 2, we ignored the factor structure of the ASC and instead applied a supervised machine learning analysis to select items that were most reliably associated with NOT signals. Multicollinearity between ASC items was expected, so LASSO regularization (Tibshirani, 1996) was performed as a method to exclude variables that did not significantly improve predictive value and which covaried strongly with reliable predictors. The logistic LASSO regression shrinks (penalizes) beta coefficients toward zero, aiming to yield conservative models that minimize overfitting. The magnitude of a penalized coefficient indicates the weight of its predictive signal, so variables with coefficients that were shrunk to exactly zero were excluded. In order to determine the optimal variable selection and model with minimal expected prediction error, a 10-fold cross-validation approach was applied (Rodriguez, Perez, & Lozano, 2010). This analysis was run across the whole dataset (weekly assessments = 895; within 283 patients) in order to identify ASC items that predicted NOT signals consistently across different partitions (folds) of the dataset. This analysis only included data from treatment week 2 onward, because the OQ-Analyst system only starts to classify sessions as on track or NOT after the initial therapy session.

In Stage 3, we examined the predictive value of alternative ASC models. Logistic multilevel modeling was applied, with weekly assessments (Level 1) nested within patients (Level 2), including random intercepts for patients. The dependent variable was the NOT signal (0 = on track; 1 = NOT). Preliminary model building steps indicated that a log linear trend for the total number of therapy sessions fit the data better than linear or exponential (quadratic and cubic) trends. The predictors entered into "Model A" included each of the four ASC domain scores derived from the original 40-item version, controlling for total number of sessions and baseline OQ-45 scores. In "Model B," again, the four ASC domain scores were entered, controlling for the total number of sessions and baseline OQ-45. However, the domain scores were derived from a shortened 30-item version of the ASC that was supported by the factor analysis (based on factor loadings >0.50). "Model C" only included the items selected by the LASSO procedure, controlling for the number of sessions and baseline OQ-45. All continuous variables were grand mean centered to aid interpretability. In order to compare these three models, we examined their goodness-of-fit (−2 log likelihood, Akaike's information criterion, and Bayesian information criterion) and predictive accuracy indices (positive and negative predictive values and area under the curve).

3 | RESULTS

3.1 | Factor analysis and reliability of the ASC

The suitability of this dataset for factor analysis was confirmed by Bartlett's test of sphericity (χ^2 = 5050.8, *df* = 780, $p < .001$) and a high index of sampling adequacy (KMO = 0.84). Table 1 presents the rotated factor loadings for each of the four ASC domains. This factor solution had an excellent goodness-of-fit index (GFI = 0.98) and produced a shortened 30-item version with reliability indices (Cronbach's α = 0.72 to 0.89) that were highly consistent with those obtained for the 40-item version. Correlations between the full and shortened scales were strong ($r > .90$, $p < .001$, shown in Table 1) for all four factors.

3.2 | LASSO variable selection

In total, 43 cases were classified as NOT at some point during therapy, and 9.9% (89/895) of all treatment weeks in the dataset showed NOT signals. The LASSO variable selection procedure produced a sparse and statistically significant model, $F(2, 892) = 16.48$, $p < .001$, that predicted the probability of showing NOT signals across treatment sessions. Of the 40 items entered as potential predictors, only two were selected: Question 14 ("I could talk about problems with my family"—social support domain), penalized $B = -0.05$, $SE = 0.03$; Question 33 ("I felt rejected or betrayed by someone"—life events domain), penalized $B = -0.04$, $SE = 0.03$.

3.3 | Comparing alternative ASC models

Table 2 presents fixed effects for logistic multilevel models examining the results of three alternative ASC models (see Section 2.4, Model A, Model B, and Model C). After controlling for treatment duration, the therapeutic alliance was not significantly associated with NOT signals in Models A and B. Motivation was significantly associated with NOT cases in Model B, but not in Model A. All three models converged in finding that greater problems (lower than average scores) in social support and life events domains were significantly associated with NOT signals. The probability of sessions being classed as NOT was not significantly associated with treatment duration ("Sessions" variable). Cases with higher intake severity on the OQ-45 measure ("BL_OQ-45" variable) were less likely to have treatment sessions classed as NOT, although this association was not statistically significant in the best fitting model (Model C).

As shown in Table 3, Model C had the best (smallest) goodness-of-fit indices, and Model B had better fit than Model A. Predictive accuracy indices were best for Model B, with an excellent trade-off between sensitivity and specificity (area under the curve = 0.98). Despite the fact that the two-item Model C was much briefer than the others, a patient with a positive test result would have an 82% probability of being classed as NOT, and a negative test result would be associated with a 94% probability of a typical symptom trajectory

TABLE 1 Factor structure of the ASC

Item	Domain	Factor 1 Explained variance = 6.269 Full 11-item α = .89 Short 10-item α = .89 Full * Short r = .99	Factor 2 Explained variance = 3.767 Full 11-item α = .77 Short 6-item α = .82 Full * Short r = .94	Factor 3 Explained variance = 5.451 Full 9-item α = .78 Short 7-item α = .81 Full * Short r = .95	Factor 4 Explained variance = 3.890 Full 9-item α = .72 Short 7-item α = .72 Full * Short r = .96
1	Alliance	0.614**	0.032	0.217	−0.028
2	Alliance	0.800**	−0.028	0.144	−0.066
3	Alliance	0.694**	0.022	0.110	−0.062
4	Alliance	0.882**	−0.023	0.021	−0.026
5	Alliance	0.686**	0.019	0.137	−0.064
6	Alliance	0.629**	−0.067	0.208	0.138
7	Alliance	0.696**	−0.082	0.112	0.222
8	Alliance	0.588**	−0.106	0.136	0.252
9	Alliance	0.471	0.240	0.116	−0.170
10	Alliance	0.868**	0.052	−0.044	−0.096
11	Alliance	0.716**	0.080	0.013	−0.129
12	Social support	−0.127	0.770**	0.145	−0.138
13	Social support	−0.218	0.832**	0.177	−0.154
14	Social support	0.127	0.529**	−0.283	0.293
15	Social support	0.100	0.591**	−0.210	0.200
16	Social support	0.074	0.703**	−0.005	0.112
17	Social support	−0.016	0.759**	−0.039	0.080
18	Social support	−0.027	0.478	0.019	−0.082
19	Social support	0.143	0.215	−0.182	−0.271
20	Social support	0.124	0.459	0.235	−0.066
21	Social support	0.338	0.169	−0.238	−0.058
22	Social support	−0.067	0.247	0.102	0.259
23	Motivation	0.230	−0.090	0.713**	−0.031
24	Motivation	0.014	−0.076	0.906**	−0.080
25	Motivation	0.203	−0.028	0.591**	−0.119
26	Motivation	−0.021	0.012	0.840**	0.003
27	Motivation	−0.099	−0.006	0.875**	−0.005
28	Motivation	0.053	0.095	0.647**	0.226
29	Motivation	−0.190	0.111	0.464	0.186
30	Motivation	−0.019	0.208	0.446	−0.139
31	Motivation	−0.042	−0.070	0.642**	0.203
32	Life events	−0.010	−0.046	−0.058	0.604**
33	Life events	0.027	0.016	−0.066	0.718**
34	Life events	0.115	−0.007	−0.030	0.694**
35	Life events	−0.054	−0.046	0.103	0.696**
36	Life events	0.032	−0.073	0.196	0.569**
37	Life events	0.002	0.081	−0.091	0.651**
38	Life events	−0.027	0.022	0.019	0.185
39	Life events	0.012	0.076	0.089	0.429
40	Life events	−0.086	−0.002	−0.001	0.575**

Note. Rotated factor loadings of parallel analysis using polychoric correlations, based on unweighted least squares, with promin rotation.
Full * Short r = correlation between full and shortened scale. Total variance explained by the factor solution was 53.1%.
*Factor loadings >0.50.

TABLE 2 Multilevel models fitting alternative versions of the ASC questionnaire to predict not-on-track signals

Variables	*B* (*SE*)	*p*	95% CI
Model A: 40-item ASC fixed effects			
Intercept	−4.53 (0.79)	<.001	−6.08, −2.97
Sessions (Log)	0.89 (0.49)	.070	−0.07, 1.85
BL_OQ-45 (mc)	−0.03 (0.01)	.002	−0.05, −0.01
Alliance (mc)	−0.23 (0.29)	.437	−0.80, 0.35
Social support (mc)	−1.06 (0.24)	<.001	−1.54, −0.58
Motivation (mc)	−0.59 (0.32)	.065	−1.21, 0.04
Life events (mc)	−1.30 (0.26)	<.001	−1.81, −0.79
Model B: 30-item ASC fixed effects			
Intercept	5.85 (1.66)	<.001	2.58, 9.11
Sessions (Log)	0.94 (0.50)	.060	−0.04, 1.91
BL_OQ-45 (mc)	−0.02 (0.01)	.023	−0.04, −0.01
Alliance (mc)	−0.33 (0.28)	.230	−0.87, 0.21
Social support (mc)	−0.61 (0.17)	<.001	−0.94, −0.28
Motivation (mc)	−0.64 (0.29)	.028	−1.22, −0.07
Life events (mc)	−0.96 (0.21)	<.001	−1.38, −0.54
Model C: 2-item ASC fixed effects			
Intercept	−3.42 (0.71)	<.001	−4.80, −2.03
Sessions (Log)	0.36 (0.46)	.427	−0.54, 1.27
BL_OQ-45 (mc)	−0.01 (0.01)	.407	−0.02, 0.01
Item 14 (mc)	−0.38 (0.13)	.003	−0.62, −0.13
Item 33 (mc)	−0.39 (0.11)	<.001	−0.61, −0.18

Note. Fixed effects from logistic multilevel models predicting risk signal classification (0 = on track, 1 = not on track); across *N* = 895 treatment sessions within *N* = 283 patients; entering random intercepts for patients; BL_OQ-45, baseline severity in OQ-45 measure; *B*, regression coefficient; CI, confidence interval; Log, log linear trend; mc, grand mean centered; SE, standard error.

("on track" signal). Model B had the best combination of positive (92%) and negative (95%) predictive values.

4 | DISCUSSION

This study demonstrates that problems with social support and adverse life events are consistently associated with extremely negative deviations in psychological distress and functioning. The present study extends previous findings (Probst et al., 2015; White et al., 2015) by identifying that interpersonal problems captured by specific ASC items (*not feeling able to speak about problems with family members*; *feeling rejected or betrayed*) were found to be particularly important. It is interesting to note that these aspects are external to the therapy process; whereas the aspects that are internal to the therapy process were either not (alliance) or not consistently (motivation) related to NOT signals.

These findings suggest that attention to interpersonal difficulties and close relationships may be an important focus of therapy in cases that are at risk of poor response to treatment. One possible interpretation could be that inpatients in this treatment setting could have significant personality dysfunction and interpersonal problems, which could become exacerbated and associated with increased distress. It is known that the prevalence of personality disorders is twice as high in inpatient settings by comparison with outpatient care and 10 times more common than in the general population (Lieb, Zanarini, Schmahl, Linehan, & Bohus, 2004). Hence, we cannot necessarily assume that the findings in this sample can be generalized to outpatient psychotherapy settings.

The four-factor structure of the ASC measure was confirmed using factor analysis. Furthermore, using a conventional item-reduction approach, the shorter 30-item version displayed equal or better internal consistency across domains, by comparison with the original 40-item version. This shortened version of the ASC had remarkable predictive values above 90%, which suggests that the information contained can reliably help to identify cases that are NOT. Although the full motivation to change scale in the 40-item ASC was not related to NOT signals, the shortened motivation to change scale in the 30-item ASC was significantly associated with NOT signals after excluding items which had low factor loadings. This implies that there are items in the full motivation to change scale that introduce noise and undermine its predictive value. On this basis, we can confidently recommend using the shortened 30-item version, which could be more feasible to implement in future studies and clinical practice. However, therapists should be instructed to carefully monitor two items (14, 33) related to interpersonal problems, which were identified as key predictors of NOT signals. This would help to make therapists more aware of specific interpersonal issues that influence treatment progress. It should, however, be kept in mind that therapists in clinical settings are rather reluctant to use ROM systems, even though using ROM has the potential to improve treatment outcomes for patients who are at risk of treatment failure (Lambert & Harmon, 2018). A recent review by Lewis et al. (2019) summarized

TABLE 3 Goodness of fit and predictive accuracy of alternative ASC models

	Goodness-of-fit index			Predictive accuracy index		
Model	−2 LL	AIC	BIC	+PV	−PV	AUC
Model A	5210.59	5212.59	5217.38	0.91	0.94	.973
Model B	5091.40	5093.40	5098.18	0.92	0.95	.975
Model C	4862.52	4864.52	4869.31	0.82	0.94	.976

Note. AIC, Akaike's information criterion; AUC, area under the curve; BIC, Bayesian information criterion; −2 LL, −2 log likelihood; −PV, negative predictive value; +PV, positive predictive value.

key barriers and facilitators for the successful implementation of ROM and feedback systems.

The following limitations should be considered when interpreting the present results. Some cases (~11%) had missing data in some of the ASC items. We dealt with this problem using multiple imputation, although this may have led to some inaccuracies in these cases. Another limitation concerns the sole reliance on the ASC to identify potential aspects that are associated with poor treatment response, whereas other variables (e.g., changes in medication, specific therapeutic techniques, and treatment adherence) could also be informative to assess the process of change. Furthermore, although we found that interpersonal aspects (captured in the *social support* and *life events* domains) were associated with NOT signals, we cannot draw any conclusions about cause-and-effect relations. Interpersonal problems may have led to increased distress in some cases, but it is also possible that increased distress (caused by other unmeasured factors) could have led to interpersonal difficulties.

We conclude that attending to social support and life events is of particular importance during the treatment of inpatients with common mental health problems, because these are found to be associated with poor treatment progress. Based on the results of our factor analysis, we recommend applying the more economical 30-item version of the ASC for future clinical and research purposes. Special attention should be paid to two specific ASC items capturing interpersonal problems. Currently, these two specific ASC features however cannot be considered as sufficient predictors of NOT cases. Their role in predicting patients' negative deviations in psychological distress has to be studied further in other samples and clinical settings.

CONFLICT OF INTEREST

Michael J. Lambert is part owner of OQ Measures, a company that owns and distributes the OQ-Analyst software discussed in this article.

ORCID

Thomas Probst https://orcid.org/0000-0002-6113-2133
Maria Kleinstäuber https://orcid.org/0000-0002-4453-507X

REFERENCES

Asay, T. P., & Lambert, M. J. (1999). The empirical case for the common factors in therapy: Quantitative findings. In M. A. Hubble, B. L. Duncan, & S. D. Miller (Eds.), (Hrsg.)*The heart and soul of change: What works in therapy* (pp. 23–55). Washington, DC: American Psychological Association.

Boswell, J. F., Kraus, D. R., Miller, S. D., & Lambert, M. J. (2015). Implementing routine outcome monitoring in clinical practice: Benefits, challenges, and solutions. *Psychotherapy Research*, *25*, 6–19. https://doi.org/10.1080/10503307.2013.817696

Finch, A. E., Lambert, M. J., & Schaalje, B. G. (2001). Psychotherapy quality control: The statistical generation of expected recovery curves for integration into an early warning system. *Clinical Psychology & Psychotherapy*, *8*, 231–242. https://doi.org/10.1002/cpp.286

Flückiger, C., Del Re, A. C., Wampold, B. E., & Horvath, A. O. (2018). The alliance in adult psychotherapy: A meta-analytic synthesis. *Psychotherapy*, *55*, 316–340. https://doi.org/10.1037/pst0000172

Hannan, C., Lambert, M. J., Harmon, C., Nielsen, S. L., Smart, D. W., Shimokawa, K., & Sutton, S. W. (2005). A lab test and algorithms for identifying clients at risk for treatment failure. *Journal of Clinical Psychology*, *61*, 155–163. https://doi.org/10.1002/jclp.20108

Harmon, S. C., Hawkins, E. J., Lambert, M. J., Slade, K., & Whipple, J. S. (2005). Improving outcomes for poorly responding clients: The use of clinical support tools and feedback to clients. *Journal of Clinical Psychology*, *61*, 175–185. https://doi.org/10.1002/jclp.20109

Harmon, S. C., Lambert, M. J., Smart, D. M., Hawkins, E., Nielsen, S. L., Slade, K., & Lutz, W. (2007). Enhancing outcome for potential treatment failures: Therapist-client feedback and clinical support tools. *Psychotherapy Research*, *17*, 379–392. https://doi.org/10.1080/10503300600702331

Haug, S., Puschner, B., Lambert, M. J., & Kordy, H. (2004). Veränderungsmessung in der Psychotherapie mit dem Ergebnisfragebogen (EB-45). *Zeitschrift für Differentielle Und Diagnostische Psychologie*, *25*, 141–151. https://doi.org/10.1024/0170-1789.25.3.141

Howard, K. I., Moras, K., Brill, P. L., Martinovich, Z., & Lutz, W. (1996). Evaluation of psychotherapy. Efficacy, effectiveness, and patient progress. *American Psychologist*, *51*, 1059–1064. https://doi.org/10.1037/0003-066X.51.10.1059

Huber, D., Brandl, T., Henrich, G., & Klug, G. (2002). Ambulant oder stationär? Eine Feldstudie zur Praxis der Indikationsstellung. *Psychotherapeut*, *47*, 16–23. https://doi.org/10.1007/s00278-001-0195-4

Lambert, M. J. (2012). Helping clinicians to use and learn from research-based systems: The OQ-Analyst. *Psychotherapy*, *49*, 109–114. https://doi.org/10.1037/a0027110

Lambert, M. J. (2013a). The efficacy and effectiveness of psychotherapy. In M. J. Lambert (Ed.), (Hrsg.)*Bergin & Garfield's handbook of psychotherapy and behavior change* (6th ed.) (pp. 169–218). New York: Wiley.

Lambert, M. J. (2013b). Outcome in psychotherapy: The past and important advances. *Psychotherapy*, *50*, 42–51. https://doi.org/10.1037/a0030682

Lambert, M. J. (2015). Progress feedback and the OQ-System: The past and the future. *Psychotherapy*, *52*, 381–390. https://doi.org/10.1037/pst0000027

Lambert, M. J., Bailey, R., Kimball, K., Shimokawa, K., Harmon, S. C., & Slade, K. (2007). *Clinical support tools manual-brief Version-40*. Salt Lake City: OQ Measures.

Lambert, M. J., Hannöver, W., Nisslmüller, K., Richard, M., & Kordy, H. (2002). Fragebogen zum Ergebnis von Psychotherapie. Zur Reliabilität und Validität der deutschen Übersetzung des Outcome Questionnaire 45.2 (OQ-45.2). *Zeitschrift für Klinische Psychologie und Psychotherapie*, *31*, 40–46. https://doi.org/10.1026/0084-5345.31.1.40

Lambert, M. J., & Harmon, K. L. (2018). The merits of implementing routine outcome monitoring in clinical practice. *Clinical Psychology: Science and Practice*, *25*, e12268.

Lambert, M. J., Morton, J. J., Hatfield, D., Harmon, S. C., Hamilton, S., Reid, R. C., & Burlingame, G. M. (2004). *Administration and scoring manual for the Outcome Questionnaire-45*. Orem, UT: American Professional Credentialing Services.

Lambert, M. J., Whipple, J. L., Bishop, M. J., Vermeersch, D. A., Gray, G. V., & Finch, A. E. (2002). Comparison of empirically derived and rationally derived methods for identifying clients at risk for treatment failure. *Clinical Psychology & Psychotherapy*, *9*, 149–164. https://doi.org/10.1002/cpp.333

Lambert, M. J., Whipple, J. L., & Kleinstauber, M. (2018). Collecting and delivering progress feedback: A meta-analysis of routine outcome monitoring. *Psychotherapy*, *55*, 520–537. https://doi.org/10.1037/pst0000167

Lewis, C. C., Boyd, M., Puspitasari, A., Navarro, E., Howard, J., Kassab, H., … Kroenke, K. (2019). Implementing measurement-based care in

behavioral health: A review. *JAMA Psychiatry*, *76*, 324–335. https://doi.org/10.1001/jamapsychiatry.2018.3329

Lieb, K., Zanarini, M. C., Schmahl, C., Linehan, M. M., & Bohus, M. (2004). Borderline personality disorder. *The Lancet*, *364*, 453–461. https://doi.org/10.1016/S0140-6736(04)16770-6

Lorenzo-Seva, U. (1999). Promin: a method for oblique factor rotation. *Multivariate Behavioral Research*, *34*, 347–356. https://doi.org/10.1207/S15327906MBR3403_3

Lutz, W., Lambert, M. J., Harmon, S. C., Tschitsaz, A., Schürch, E., & Stulz, N. (2006). The probability of treatment success, failure and duration-what can be learned from empirical data to support decision making in clinical practice? *Clinical Psychology & Psychotherapy*, *13*, 223–232. https://doi.org/10.1002/cpp.496

Norcross, J. C., Krebs, P. M., & Prochaska, J. O. (2011). Stages of change. *Journal of Clinical Psychology*, *67*, 143–154. https://doi.org/10.1002/jclp.20758

Probst, T., Lambert, M. J., Dahlbender, R. W., Loew, T., & Tritt, K. (2014). Providing patient progress feedback and clinical support tools to therapists: Is the therapeutic process of patients on-track to recovery enhanced in psychosomatic in-patient therapy under the conditions of routine practice? *Journal of Psychosomatic Research*, *76*, 477–484. https://doi.org/10.1016/j.jpsychores.2014.03.010

Probst, T., Lambert, M. J., Loew, T., Dahlbender, R. W., Göllner, R., & Tritt, K. (2013). Feedback on patient progress and clinical support tools for therapists: Improved outcome for patients at risk of treatment failure in psychosomatic in-patient therapy under the conditions of routine practice. *Journal of Psychosomatic Research*, *75*, 255–261. https://doi.org/10.1016/j.jpsychores.2013.07.003

Probst, T., Lambert, M. J., Loew, T., Dahlbender, R. W., & Tritt, K. (2015). Extreme deviations from expected recovery curves and their associations with therapeutic alliance, social support, motivation, and life events in psychosomatic in-patient therapy. *Psychotherapy Research*, *25*, 714–723. https://doi.org/10.1080/10503307.2014.981682

Rodriguez, J. D., Perez, A., & Lozano, J. A. (2010). Sensitivity analysis of k-fold cross validation in prediction error estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*, 569–575. https://doi.org/10.1109/TPAMI.2009.187

Roehrle, B., & Strouse, J. (2008). Influence of social support on success of therapeutic interventions: A meta-analytic review. *Psychotherapy*, *45*, 464–476. https://doi.org/10.1037/a0014333

Schafer, J. L., & Olsen, M. K. (1998). Multiple imputation for multivariate missing-data problems: A data analyst's perspective. *Multivariate Behavioral Research*, *33*, 545–571. https://doi.org/10.1207/s15327906mbr3304_5

Sexton, H. (1996). Process, life events, and symptomatic change in brief eclectic psychotherapy. *Journal of Consulting and Clinical Psychology*, *64*, 1358–1365. https://doi.org/10.1037/0022-006X.64.6.1358

Shimokawa, K., Lambert, M. J., & Smart, D. W. (2010). 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*, *78*, 298–311. https://doi.org/10.1037/a0019247

Slade, K., Lambert, M. J., Harmon, S. C., Smart, D. W., & Bailey, R. (2008). Improving psychotherapy outcome: The use of immediate electronic feedback and revised clinical support tools. *Clinical Psychology & Psychotherapy*, *15*, 287–303. https://doi.org/10.1002/cpp.594

Spielmans, G. I., Masters, K. S., & Lambert, M. J. (2006). A comparison of rational versus empirical methods in prediction of negative psychotherapy outcome. *Clinical Psychology & Psychotherapy*, *13*, 202–214. https://doi.org/10.1002/cpp.491

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B: Methodological*, *58*, 267–288.

Timmerman, M. E., & Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods*, *16*(2), 209–220. https://doi.org/10.1037/a0023353

Warren, J. S., Nelson, P. L., Burlingame, G. M., & Mondragon, S. A. (2012). Predicting patient deterioration in youth mental health services: Community mental health vs. managed care settings. *Journal of Clinical Psychology*, *68*, 24–40. https://doi.org/10.1002/jclp.20831

Whipple, J. L., Lambert, M. J., Vermeersch, D. A., Smart, D. W., Nielsen, S. L., Hawkins, E. J., & Eric, J. (2003). Improving the effects of psychotherapy: The use of early identification of treatment and problem-solving strategies in routine practice. *Journal of Counseling Psychology*, *50*, 59–68. https://doi.org/10.1037/0022-0167.50.1.59

White, M. M., Lambert, M. J., Ogles, B. M., Mclaughlin, S. B., Bailey, R. J., & Tingey, K. M. (2015). Using the Assessment for Signal Clients as a feedback tool for reducing treatment failure. *Psychotherapy Research*, *25*, 724–734. https://doi.org/10.1080/10503307.2015.1009862

Zipfel, S., Herzog, W., Kruse, J., & Henningsen, P. (2016). Psychosomatic medicine in Germany: More timely than ever. *Psychotherapy and Psychosomatics*, *85*, 262–269. https://doi.org/10.1159/000447701

How to cite this article: Probst T, Kleinstäuber M, Lambert MJ, et al. Why are some cases not on track? An item analysis of the Assessment for Signal Cases during inpatient psychotherapy. *Clin Psychol Psychother*. 2020;27:559–566. https://doi.org/10.1002/cpp.2441