Psychometric Validation of the Treatment Effectiveness Assessment in Patients with Opioid Use Disorder

Walter Ling[1] | David Farabee[1] | Brian Perrochet[2] | Vijay R. Nadipelli[3] | Naoko A. Ronquest[3] | Caitlyn T. Solem[4] | Chinmay Deshpande[4]

[1] UCLA Integrated Substance Abuse Programs, Los Angeles, California, USA. [2] Marron Institute, New York University, New York, New York, USA. [3] Indivior Inc., Richmond, Virginia, USA.
[4] Pharmerit International, Bethesda, Maryland, USA.

Introduction and Aims

- Patient experience is starting to play a larger role in regulatory decision making in opioid use disorder (OUD)[1]; currently, there is no validated instrument that can be easily incorporated into clinical practice and research.
- The Treatment Effectiveness Assessment (TEA) is an efficient, patient-centered instrument for evaluating progress in recovery from OUD.[2,3]
- The TEA consists of 4 single-item domains (Substance Use, Health, Lifestyle, and Community) measuring treatment-related improvement. Domain scores range from 1 ("not much") to 10 ("much better") and total scores range from 4 to 40.
- Since its publication in 2012, the TEA has been adopted in clinical practice and research.
- This study evaluated the TEA's psychometric properties among treated participants with moderate to severe OUD and determined minimally important differences (MIDs) for total and subscale scores.

Design and Methods

Study Design

- This analysis was based on data in a Phase-III open-label safety study of BUP-XR (buprenorphine extended-release monthly injection, for subcutaneous use [CIII], SUBLOCADE™), known as RBP-6000 during development, among participants with moderate to severe OUD (NCT02510014).[4]
- Participants completed the TEA and other measures (**Figure 1**). Baseline was defined as the measurements at injection 1.
- Only participants who answered at least 1 TEA item at baseline were included in the present psychometric validation analysis.

Statistical Analysis

- Internal consistency for TEA total score was assessed at baseline using Cronbach's alpha (threshold: 0.7).[5]
- Test-retest reliability (threshold: 0.7) for TEA single-item domains and total scores was examined among participants (n=177) with stable urine drug screen results between injection 12 and end of study (EOS).[6, 7]
- Known groups validity was assessed using 1-way ANOVA, grouping patients by current health status (SF-36v2 Item 1).
- Convergent/divergent validity were evaluated using Pearson's correlation coefficient to test for correlations between TEA domain and total scores with baseline EuroQol-5D-5L (EQ-5D-5L) index and visual analog scale (VAS) scores; 36-item Short Form Health Survey, version 2 (SF-36v2) mental and physical component scores; Addiction Severity Index Lite (ASI-Lite) dimensions; and opioid craving VAS (OC VAS).
- Ability to detect change was evaluated by comparing change in TEA scores to change in SF-36v2 Item 1 (current health status) from screening to baseline.
- MID estimates were evaluated using distributional methods (i.e., calculating 0.5 standard deviation at baseline and standard error of measurement, computed as SD x √ [1-reliability]); anchor-based methods (change from baseline among participants in whom ≥80% of weekly urine drug screen and self-report results were negative between injection 2 and EOS)[8-10]; and receiver operating characteristic (ROC) curves (to assess what change in TEA total score from baseline to EOS would most discriminate between participants who achieved the ≥80% threshold of negative opioid use).[11]

Figure 1. Schedule of Assessments (Select Patient-reported Measures)

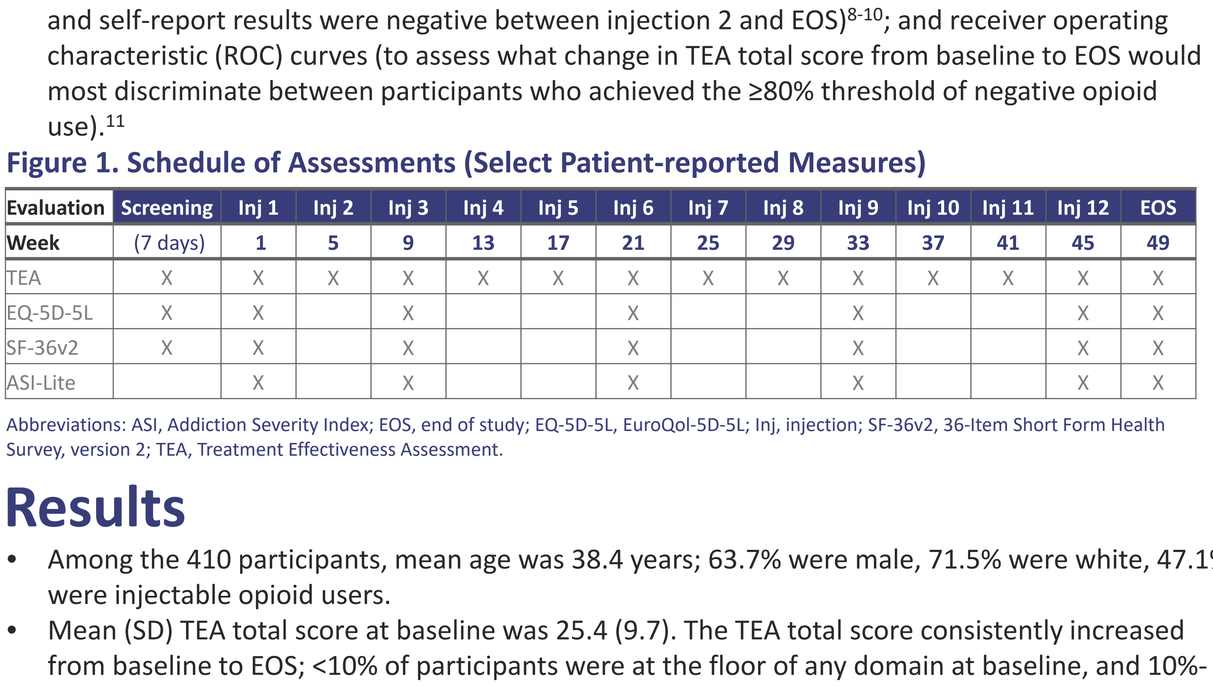
Evaluation	Screening	Inj 1	Inj 2	Inj 3	Inj 4	Inj 5	Inj 6	Inj 7	Inj 8	Inj 9	Inj 10	Inj 11	Inj 12	EOS
Week	(7 days)	**1**	**5**	**9**	**13**	**17**	**21**	**25**	**29**	**33**	**37**	**41**	**45**	**49**
TEA	X	X	X	X	X	X	X	X	X	X	X	X	X	X
EQ-5D-5L	X	X		X			X			X			X	X
SF-36v2	X	X		X			X			X			X	X
ASI-Lite		X		X			X			X			X	X

Abbreviations: ASI, Addiction Severity Index; EOS, end of study; EQ-5D-5L, EuroQol-5D-5L; Inj, injection; SF-36v2, 36-Item Short Form Health Survey, version 2; TEA, Treatment Effectiveness Assessment.

Results

- Among the 410 participants, mean age was 38.4 years; 63.7% were male, 71.5% were white, 47.1% were injectable opioid users.
- Mean (SD) TEA total score at baseline was 25.4 (9.7). The TEA total score consistently increased from baseline to EOS; <10% of participants were at the floor of any domain at baseline, and 10%-20% were at the ceiling across domains (**Figure 2**).

Figure 2. Item Distribution and Floor-Ceiling Effects in TEA Analysis Population at Baseline (N=410)



Abbreviations: TEA, Treatment Effectiveness Assessment.
[a] An item may exhibit floor or ceiling effects if more than (100/the number of response options)% of responses are in the highest or lowest response categories.

Results, continued

- The TEA exhibited strong internal consistency (Cronbach's α = 0.897), which was retained when each item was removed and internal consistency was re-calculated. When considering test-retest reliability, the intraclass correlation for the total score approached 0.7 (intraclass correlation coefficient [ICC]: 0.688; 95% CI: 0.602-0.757; **Table 1**).
- The TEA total score demonstrated differentiation between known groups by current health status (P<.001). For convergent and divergent validity, the directionality of correlations of TEA scores with other outcomes were consistent with hypotheses, but weaker than anticipated between like domains (**Table 1**).
- The TEA instrument was sensitive to change in current health status (**Table 2**).
- MIDs for TEA total score were highly dependent on method used. When using distribution-based measures and ROC methods, MIDs ranged from 4.8 to 8.0. However, anchoring results on having ≥80% urine drug screens negative between week 5 and EOS, the mean MID was 6.8 (**Table 3**).

Table 1. Measures of TEA Reliability and Validity

	N	TEA Domain				TEA Total Score
		Substance Use	Health	Lifestyle	Community	
Measures of Reliability						
ICC (95% CI)[a]	177	0.647 (0.554-0.724)	0.708 (0.627-0.774)	0.649 (0.556-0.726)	0.609 (0.508-0.693)	0.688 (0.602-0.757)
Cronbach's alpha	410	0.927[b]	0.843[b]	0.837[b]	0.857[b]	0.897
Known Groups Validity						
Current health status (SF-36v2 Item 1), no. (%)						
Excellent	56	7.3 (2.9)	7.1 (2.9)	7.7 (2.6)	7.4 (2.8)	29.4 (9.8)
Very Good	167	6.5 (3.1)	6.5 (2.7)	6.7 (2.6)	6.8 (2.7)	26.5 (9.7)
Good	140	6.3 (3.0)	5.6 (2.5)	6.1 (2.4)	6.1 (2.5)	24.0 (9.1)
Fair/Poor	47	5.9 (2.9)	4.5 (2.5)	5.3 (2.6)	5.3 (3.1)	21.1 (9.0)
P-value[c]		.097	<.001	<.001	<.001	<.001
Convergent Validity						
OC VAS		-0.151	-0.14	-0.113	-0.133	-0.066
ASI medical status		-0.088	-0.065	-0.003	-0.044	-0.003
ASI employment status		-0.027	-0.014	-0.040	-0.053	-0.097
ASI alcohol use		-0.065	0.001	0.040	0.003	0.032
ASI drug use		-0.109	-0.076	-0.084	-0.115	-0.123
ASI legal status		-0.055	-0.082	-0.057	-0.084	-0.096
ASI social/family status		-0.085	-0.057	-0.041	-0.052	-0.003
ASI psychiatric status		-0.008	0.067	0.083	0.06	0.063
SF-36v2 PCS		0.161	0.114	0.065	0.106	0.037
SF-36v2 MCS		0.296	0.263	0.264	0.247	0.058
EQ-5D Utility Index		0.341	0.278	0.257	0.301	0.184
EQ-5D VAS		0.203	0.157	0.173	0.165	0.054

Abbreviations: ASI, Addiction Severity Index; EQ-5D, EuroQol-5D; ICC, intraclass correlation coefficient; MCS, mental component score; OC, opioid craving; PCS, physical component score; SF-36v2, SF-36 version 2; VAS, visual analogue scale.
[a] Stable participants were defined based on last 2 visits in the study and limited to those with stable urine drug screen results between these last visits.
[b] Alpha if item deleted.
[c] P-value is based on the overall ANOVA.

Table 2. Ability to Detect Change for TEA

Current Health Status from Screening to Baseline Based on SF-36v2 Item 1 (current health status)[a]	Change in TEA Total Score From Screening to Baseline, Mean (SD)
Improved (n=189)	13.5 (10.9)
No change (n=177)	11.5 (10.5)
Decline (n=44)	8.7 (9.7)

Abbreviations: SF-36v2, 36-Item Short Form Survey version 2; TEA, Treatment Effectiveness Assessment.
[a] Participants were grouped as those who experienced improvement of ≥1 point, no change, or declined ≥1 point between screening and baseline.

Table 3. Summary of MID Estimates for TEA Total Score Using TEA Analysis Population

Domain	0.5 SD	SEM	Anchored (UDS)	ROC
TEA Total Score	4.8	5.6	6.8	8.0

Abbreviations: MID, minimally important difference; ROC, receiver operating characteristic curve; TEA, Treatment Effectiveness Assessment; UDS, urine drug screen.
[a] SEM is calculated as [SD x √(1-reliability)] at baseline; 0.5 SD calculated as 1/2 standard deviation at baseline.

Discussion and Conclusions

- The TEA is an instrument that is easy to administer and can help clinicians and patients benchmark and assess recovery in clinical research as well as in real-world treatment settings.
- The TEA demonstrated evidence of reliability and validity in this moderate to severe OUD cohort.
- This study provides a range of MIDs to further inform definitions of meaningful change in TEA scores.
- Further research is needed to evaluate the reliability, validity, and responsiveness of TEA in other clinical settings and to identify recovery-related outcomes that are strongly correlated with TEA scores.

References

1. US Food and Drug Administration. https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm615892.htm. August 2018.
2. Ling W, Farabee D, Liepa D, Wu LT. The Treatment Effectiveness Assessment (TEA): an efficient, patient-centered instrument for evaluating progress in recovery from addiction. *Subst Abuse Rehabil*. 2012;3:129.
3. Betty Ford Institute Consensus Panel. What is recovery? A working definition from the Betty Ford Institute. *J Subst Abuse Treat*. 2007;33(3):221.
4. Ling W, Nadipelli VR, Solem CT, et al. Impact of RBP-6000 (once-monthly depot buprenorphine) on patient reported outcomes: a long-term study. American Society of Addiction Medicine 49th Annual Conference (2018). San Diego, CA.
5. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16(3):297-334.
6. Oliva TA, Oliver RL, MacMillan IC, et al. A catastrophe model for developing service satisfaction strategies. *J Marketing*. 1992;56(3):83-95. doi:10.2307/1252298
7. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86(2):420-428.
8. Jaeschke R, Singer J, Guyatt GH, et al. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials*. 1989;10(4):407-415.
9. Norman GR, Sloan JA, Wyrwich KW, et al. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care*. 2003;41(5):582-592. doi:10.1097/01.Mlr.0000062554.74615.4c
10. Revicki D, Hays RD, Cella D, et al. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol*. 2008;61(2):102-109. doi:10.1016/j.jclinepi.2007.03.012
11. Froud R, Abel G. Using ROC curves to choose minimally important change thresholds when sensitivity and specificity are valued equally: the forgotten lesson of Pythagoras. Theoretical considerations and an example application of change in health status. *PLoS One*. 2014;9.12 (2014): e114468.

Disclosures

WL is a consultant to Indivior Inc., Alkermes/Braeburn, Opiant, and Titan Pharma. DF is a consultant to Indivior Inc. DF has received study medications from Alkermes. BP is an independent consultant and declares nothing to disclose nor conflicts of interest. VRN and NAR are employees of Indivior Inc. CTS and CD are employees of Pharmerit International and are consultants to Indivior Inc. The study was funded by Indivior Inc.

Contact Information: Walter Ling at lwalter@ucla.edu