Research & Policy Brief

Measure and Data Element Identification for the HRSA Evidence-Based Tele-Behavioral Health Network Program and the HRSA Substance Abuse Treatment Telehealth Network Grant Program

Marcia M. Ward, PhD;[A] Kimberley Fox, MPA;[B] Kimberly Merchant, MA;[A] Amanda Burgess, MPPM;[B] Fred Ullrich, BA;[A] Lily Parenteau, MPH;[B] Jean Talbot, PhD, MPH;[B] Karen Pearson, MLIS, MA;[B] Jennifer Mackenzie, BA;[B] Whitney Wiggins, MA;[C] Stephanie Miller, BS[C]

[A]University of Iowa, College of Public Health, Iowa City, IA; [B]University of Southern Maine, Portland, ME; [C]HRSA FORHP Office for the Advancement of Telehealth, Rockville, MD

Introduction and Background

Behavioral health (mental health and substance use disorder) needs are significant and largely remain unmet in rural areas.[1] For example, 113 million people reside in Mental Health Professional Shortage Areas, mostly in rural areas of the U.S.[2] A specific concern for rural communities is the spiraling substance use crisis.[3,4] In 2015, only 10.8 percent of the 21.7 million people who required substance use disorder treatment received those services.[5] Absence of social support, stigma associated with substance use disorder, privacy concerns, time conflicts, lack of availability of appropriate treatment options, and difficulty in accessing treatment are some of the individual, contextual, and systemic barriers that obstruct treatment for substance use disorder.[6] These issues along with the existing provider shortages make rural care delivery for behavioral health uniquely challenging. Telehealth is seen as one possible solution for expanding behavioral health treatment services and making them accessible to those who are not utilizing these services. Attributes like the flexible nature of services, improved accessibility, ability to reach those who cannot be reached through traditional counseling services, and reducing stigma associated with seeking care directly from a specialty center make telehealth services an appealing and feasible alternative.[7]

Key Findings

- Candidate measures were identified through an extensive published and grey literature review and search of key organizations' existing measures.

- The resulting inventory was subjected to three rounds of scoring by the research team, external experts, and grantees.

- The final set of 9 measures were operationalized into 26 data elements needed for systematic data collection and statistical analysis.

- The Evidence-Based Tele-Behavioral Health Network Program (EB THNP) grantees and Substance Abuse Treatment Telehealth Network Grant Program (SAT TNGP) grantees are reporting data on services they deliver.

To promote this potential, the Federal Office of Rural Health Policy (FORHP) in the Health Resources and Services Administration (HRSA) funded two grant programs to help provide telehealth services to rural populations with mental health and substance use disorder needs. The Evidence-Based Tele-Behavioral Health Network Program (EB THNP)[a] funded 14 grantees, and the Substance Abuse Treatment Telehealth Network Grant Program (SAT TNGP)[b] funded 3 grantees. Concurrently, the Rural Telehealth Research Center (RTRC)[c] worked cooperatively with FORHP to

establish data collection protocols and tools on a set of measures that could be used for a cross-grantee evaluation of tele-behavioral health services.

Measure Scoring Criteria and Review Process
This project began with a review of published literature for studies related to tele-behavioral health, with a particular focus on systematic reviews and meta-analyses. The search yielded 101 reviews, from which we identified 1,871 unique studies. Based on guidance provided by FORHP related to their priority areas of focus, we identified four broad measure domains by which we categorized measures—clinical outcomes, cost and cost-effectiveness, quality of care, and access. To narrow the 1,871 unique studies, we first focused only on studies that were cited in more than 1 systematic review, which resulted in 477 studies. The research team reviewed the articles and extracted 1,304 relevant measures, recording them in a spreadsheet.

The research team also conducted an environmental scan of performance and outcome measures developed or recommended by stakeholders. The team developed a list of relevant stakeholder organizations, yielding 36 agencies and organizations for review. A researcher conducted comprehensive reviews of each agency's website and included searches for specific terms: tele-behavioral health, telepsychiatry, telehealth, telemedicine, measure, and evaluation. The environmental scan produced 106 measures. To evaluate the combined list of candidate measures, separate sets of scoring criteria were developed for the clinical and nonclinical measures.

Nonclinical scoring process (Rounds 1 & 2)
Prior to scoring, 145 patient satisfaction measures were removed from the database of potential measures based on a literature review that revealed that high levels of patient satisfaction with telehealth have already been reported across numerous studies. For Round 1 of nonclinical scoring, each of the remaining 527 nonclinical measures were scored by 6 team members on 2 criteria (measure specification and utility for evaluation of grant objectives). Throughout, scoring was tailored to each criterion but generally used a 3-level scale from 0 = "not at all" to 2 = "high." To determine which measures to carry through to the second round of scoring, scores were compared within and between 38 subdomains (Table 1) within each of 4 thematic domains (clinical outcomes, cost and cost-effectiveness, quality of care, and access). Measures with a total score above the average in their domain were retained and moved to the second round of scoring. In addition, measures with scores above the subdomain average but below the domain average were reviewed to retain any that captured missing concepts. For Round 2, the resulting 183 nonclinical measures were reviewed by a separate set of 6 team members. Measures were scored on three criteria (amenable to telehealth, data collection feasibility, and importance to FORHP evaluative needs). A researcher compiled all scores and set a threshold using the same method as was used in Round 1. The team again reviewed measures with scores above the subdomain average but below the domain average to determine which measures to move to Round 3.

Clinical scoring process (Rounds 1 & 2)
Prior to scoring, 265 clinical measures were removed from the database of potential measures because they did not align with the target conditions proposed by grantees. For Round 1 of clinical scoring, 5 team members scored the remaining 667 measures on 2 criteria (measure specification and utility for evaluation of grant objectives). Similar to the nonclinical process, measures with a total score below average in the subdomain and domain were deleted. Within a group of measures that used the same screening tool, the measure with the highest measure-specification score was retained and the remaining measures were deleted. A total of 203 clinical measures moved on to Round 2 scoring, which was completed by a separate set of 5 team members. Measure scores from the three scoring criteria (amenable to telehealth, data collection feasibility, and importance to FORHP evaluative needs) were compiled following a similar process as the other rounds of scoring and reviewed by the team.

Feedback from content-area experts
Following the second round of scoring for both the nonclinical and clinical measures, a health economist reviewed the remaining measures in the cost and cost-effectiveness domain. Another researcher, external to the project team and with expertise in substance use disorders, reviewed nonclinical and clinical measures related to substance use. Feedback from the two experts was incorporated into the review process and used to further narrow the list of measures moving to the third round of scoring. In addition, three grantees with particular expertise in the field of

telehealth were contacted. We sought their guidance as a way to narrow the total number of measures and reduce the burden on all grantees for the third round of scoring. The 3 expert grantees were sent 172 measures to review. Rather than scoring each measure, as was done in previous rounds, the grantee experts were instructed to indicate which measures they would keep and which they would exclude from further scoring. Experts were also encouraged to provide comments explaining their choices, and to provide any other feedback they had about the measures. This feedback from experts was used to further narrow the measures list.

Nonclinical and clinical scoring by grantees (Round 3)
Before sending the measures to the grantees for Round 3 scoring, FORHP and RTRC hosted a webinar with the EB THNP grantees to provide a greater level of detail about the measure selection process and explain what was expected of them for Round 3 scoring. Following the webinar, an inventory of 69 prospective measures was sent to each grantee organization for scoring on 3 criteria (clinical applicability and relevance, data collection feasibility, and utility for intended stakeholders). Table 2 displays the steps taken to arrive at the 69 prospective measures. All 14 EB THNP grantees provided feedback on the list of measures. Scores used a similar scale as in previous rounds of scoring. Grantee scores were compiled into a single spreadsheet, where an average score for each criterion for each measure was calculated. Written comments provided by grantees were also compiled in a separate document, organized by measure.

Table 2. Number of Measures by Domain Reviewed in Each Round of Scoring

Number of measures scored in each round by domain				
Domain	**Round 1**	**Round 2**	**Expert Review**	**Round 3**
Clinical Outcomes	667*	203	91	33
Cost/Cost-effectiveness	117	38	17	8
Quality of care	260**	79	39	14
Access	150	66	25	14
Total	**1194**	**386**	**172**	**69**

*This number does not include the 265 clinical outcomes measures removed before Round 1 scoring because they did not align with proposed grantee target conditions.
**This number does not include the 145 patient satisfaction measures removed before Round 1 scoring.

Final Selection of Measures and Assessment Instruments
RTRC researchers winnowed the remaining measures through iterative reviews of average measure scores and comments collected during the grantee review process. Measures with low average scores, a high number of "0" scores from grantees, or mostly negative comments from grantees were generally eliminated. Among the clinical outcome measures, many of the measures using survey instruments that were lengthy and/or had a cost associated with their use were rated poorly by grantees. Duplicative measures were combined with similar measures or were eliminated. The result of this process was a narrowed list of 11 measures, shown in Table 3.

An important step in the measure selection process was finalizing the survey instruments for each of the clinical outcome measures. The research team reviewed prospective instruments and compiled psychometric properties and other background information about each instrument. In particular, the team sought instruments that had been used previously to track change over time, had no associated costs, and were a reasonable length for grantees to administer. The Patient Health Questionnaire-9[8] for depression and the Generalized Anxiety Disorder 7-item scale[9] for anxiety-related measures were well received by grantees during the review process and were retained. After reviewing a number of Quality of Life/Functioning instruments, the research team selected the Patient-Reported Outcomes Measurement Information System (PROMIS) Global Health scale version 1.2.[10] The PROMIS Global Health scale provides mental health and physical health component scores as well as an overall total score. All three scores can be used to track changes in patient functioning and well-being over time. After considerable review of available measures for substance use disorder, the team selected the Drug Use Disorders Identification Test-Consumption (DUDIT-C),[11] a brief instrument that is used for recent drug-related problems. Unlike many other measures that are focused on screening, the DUDIT-C has been used previously to measure change in drug use severity over time.

Data Elements Needed for Systematic Data Collection and Analysis

The measures were operationalized into a set of 26 component data elements, shown in Table 4. RTRC created a dictionary of all data elements to define terms, indicate allowable values, and provide abstractor notes. In addition to a training manual, an Excel-based tool, termed the Behavioral-Telehealth Evidence Collection Tool (B-TEC Tool), was created for data collection. Data use agreements were established between RTRC and each grantee, and all involved entities secured Institutional Review Board Human Subject Review approval. To facilitate both the signing of the data use agreements and Institutional Review Board approval, no protected health information was involved and data were deidentified prior to transmission to RTRC. RTRC performed data monitoring and management activities to verify data accuracy, completeness, consistency, and timeliness. OMB clearance was received in October 2019 and grantees will collect data until the conclusion of their grant period (August 31, 2020, for SAT TNGP and August 31, 2021, for EB THNP).

Significance

The Agency for Healthcare Research and Quality published an evidence map to synthesize the quantity and quality of studies in telehealth and found that behavioral health was the fourth leading focus of telehealth systematic reviews.[12] As telehealth use grows, efforts to examine the evidence base for telehealth applications are hampered by the disparate structure, process, and outcome measures used in telehealth research. Policies to reduce the barriers to telehealth and further its useful expansion will rely on sound evidence of its effectiveness, especially when multiple studies replicate positive findings on a common set of measures. Given the prevalence of behavioral health conditions in the U.S. and the applications of telehealth services addressing this need, efforts to define a common set of measures are timely and will contribute to establishing the evidence base examining their effectiveness.

Notes

1. Gale J, Janis J, Coburn A, Rochford H. Behavioral health in rural America: Challenges and opportunities. RUPRI Health Panel Report. December 2019.
2. HRSA. Health Professional Shortage Areas (HPSAs). 2019. Retrieved from: https://bhw.hrsa.gov/shortage-designation/hpsas
3. Keyes KM, Cerda M, Brady JE, Havens JR, Galea S. Understanding the rural-urban differences in nonmedical prescription opioid use and abuse in the United States. *Am J Public Health*. 2014;104:e52-e59.
4. Monnat SM, Rigg KK. Examining rural/urban differences in prescription opioid misuse among US adolescents. *J Rural Health*. 2016;32:204-218.
5. Bose J, Hedden SL, Lipari RN, Park-Lee E. Key substance use and mental health indicators in the United States: Results from the 2015 National Survey on Drug Use and Health. HHS Publication No. SMA 16-4984, NSDUH Series H-51. Retrieved from: http://www.samhsa.gov/data/
6. Rapp RC, Xu J, Carr CA, Lane T, Wang J, Carlson R. Treatment barriers identified by substance abusers assessed at a centralized intake unit. *J Subst Abuse Treat*. 2006;30(3):227–235. Retrieved from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1986793/
7. Quaglio G, Esposito G. Technological innovation strategies in substance use disorders. European Parliamentary Research Service Scientific Foresight Unit (STOA), 2017. Retrieved from: http://www.europarl.europa.eu/RegData/etudes/STUD/2017/581942/EPRS_STU(2017)581942_EN.pdf
8. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine.* 2001;16(9):606-13.
9. Spitzer RL, Kroenke K, Williams JB, Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine.* 2006;166(10):1092-1097.
10. Hays RD, Bjorner JB, Revicki DA, Spritzer KL. Development of physical and mental health summary scores from the Patient-Reported Outcomes Measurement Information System (PROMIS) Global Items. *Qual Life Res.* 2009;18(7):873-880.
11. Berman AH, Bergman H, Palmstierna T, Schlyter F. *DUDIT Manual (the Drug Use Disorders Identification Test).* Stockholm, Sweden: Karolinska Institutet, Department of Clinical Neuroscience; March, 2003. Version 1.0.
12. Totten AM, Womack DM, Eden KB et al. Telehealth: Mapping the evidence for patient outcomes form systematic reviews. Technical Brief No. 26 (Prepared by the Pacific Northwest Evidence-based Practice Center under Contract No. 290-2015-00009-1.) AHRQ Publication No. 16-EHC034-EF.

Table 1. Measure Domains and their Subdomains

Access	Clinical Outcomes	Cost /Cost-effectiveness	Quality of Care
Cultural access	Adherence	Broad resource utilization	Accuracy of assessment
Digital access	Functional improvement	Burden on social network	Coordination of care
Distance to service	Symptom outcomes	Cost avoidance	Integration of care
Length of session		Economic evaluation	Motivational readiness
Likelihood to access versus traditional care		Facilities and maintenance	Patient safety
Number of services		Missed obligations	Patient satisfaction
Numbers served		Personnel	Provider satisfaction
Treatment utilization		Public versus private funding	Rapport/Therapeutic alliance
Wait times		Supplies	Stigma
		Technology direct	Telehealth quality
		Technology indirect	Usability
		Training	
		Travel direct	
		Travel indirect	
		Value proposition	

Table 3. Description of 9 Measures for the EB THNP/SAT TNGP Cross-Grantee Evaluation

Domain		and Brief Description
Access	1	Number of behavioral health encounters per patient, by intervention group (telehealth vs nontelehealth): total and by service type, provider type, and disposition
	2	Percent of scheduled behavioral health encounters not completed, by intervention group (telehealth vs nontelehealth): total, by service type, and specific reason
Cost and Cost Effectiveness	3	Percent of encounters where behavioral health services are billed to insurance, by intervention group (telehealth vs nontelehealth): total and by service type
	4	Estimated reimbursement for behavioral health services, by intervention group (telehealth vs nontelehealth): total and by service type
	5	Estimated travel costs for behavioral health services, by intervention group (telehealth vs nontelehealth): total and by service type
Clinical Outcomes	6	Change in mental and physical health quality of life over time, by intervention group (telehealth vs nontelehealth): total and by service type. Measure will use the Patient-Reported Outcomes Measurement Information System (PROMIS) Global Health to assess patient functioning.
	7	Change in depression symptoms over time, by intervention group (telehealth vs nontelehealth): total and by service type. Measure will use the Patient Health Questionnaire-9 (PHQ-9) to assess depression symptoms in patients whose primary complaint is depression or who have a primary or secondary ICD-10 indicative of depression.
	8	Change in generalized anxiety symptoms over time, by intervention group (telehealth vs nontelehealth): total and by service type. Measure will use the Generalized Anxiety Disorder Scale-7 (GAD-7) to assess anxiety symptoms in patients whose primary complaint is anxiety or who have a primary or secondary ICD-10 indicative of anxiety.
	9	Change in substance use severity over time, by intervention group (telehealth vs non-telehealth): total and by service type. Measure will use the Drug Use Disorders Identification Test – Consumption (DUDIT-C) to assess substance use severity in patients whose primary complaint is substance use or who have a primary or secondary ICD-10 indicative of substance use.

Table 4. Description of 26 Data Elements for the EB THNP/SAT TNGP Cross-Grantee Evaluation

Level	Data Elements
Patient	Treatment group – Indicates whether the patient was in the telehealth group or the nontelehealth comparison group
Patient	Age – The patient's age at intake
Patient	Sex – The patient's sex
Patient	Race – The patient's racial group
Patient	Ethnicity – The patient's ethnic group
Patient	Patient's insurance status – The type of insurance that the patient has at intake
Patient	Patient travel miles to the initial planned place of behavioral health services – Miles from the patient's location to where the patient plans to receive behavioral health services
Patient	Patient travel time to the initial planned place of behavioral health services – Travel time from the patient's location to where the patient plans to receive behavioral health services
Patient	Patient travel miles to next likely source of behavioral health services – Miles from the patient's location to the next likely source of behavioral health services if the planned place of services was not available
Patient	Patient travel time to next likely source of behavioral health services – Travel time from the patient's location to the next likely source of behavioral health services if the planned place of services was not available
Patient	Patient likelihood of using next source of behavioral health services – The patient's likelihood of using next source of care for type of service delivered
Intake + Repeat	Assessment instrument administration timing – The number of weeks since the initiation of the treatment when the assessment instrument(s) were re-administered
Intake + Repeat	PROMIS Global Health – Mental Health score (component)
Intake + Repeat	PROMIS Global Health – Physical Health score (component)
Intake + Repeat	PROMIS Global Health score (total)
Intake + Repeat	PHQ-9 depression symptoms score
Intake + Repeat	GAD-7 generalized anxiety symptoms score
Intake + Repeat	DUDIT-C substance use severity score
Encounter	Treatment type – Whether encounter was planned for telehealth or nontelehealth services
Encounter	Timing of encounter – Number of days since first treatment encounter
Encounter	Therapy scheduling success – Whether scheduled session was completed
Encounter	Provider type – Type of provider/clinician seen for behavioral health services during this encounter
Encounter	Patient's behavioral health diagnosis – The ICD-10 code(s) associated with the diagnosis established to be chiefly responsible for the behavioral health services
Encounter	Treatment service type – CPT code for each encounter
Encounter	Disposition recommendation – Indicates the provider's recommended disposition for the patient at the end of the encounter
Encounter	Treatment billing – Indicates whether the behavioral health services encounter was billed to insurance

RHRC
Rural Health Research & Policy Centers
Funded by the Federal Office of Rural Health Policy
www.ruralhealthresearch.org