AMERICAN
TELEMEDICINE
ASSOCIATION

A LEXICON OF ASSESSMENT AND OUTCOME MEASURES FOR TELEMENTAL HEALTH

NOVEMBER 2013

A Lexicon of Assessment and Outcome Measures for Telemental health

November 2013

Jay H. Shore MD, MPH[1], Matt C. Mishkind, PhD[2], Jordana Bernard, MBA[3], Charles R. Doarn, MBA[4], Iverson Bell, Jr., MD[5], Rajiv Bhatla, MD, FRCPC, DABN[6], Elizabeth Brooks, PhD[1], Robert Caudill, MD[7], Ellen R. Cohn, PhD, CCC-SLP[8], Barthold J. Delphin[9], Antonio Eppolito, MD[10], John Fortney, PhD[11], Karl Friedl, PhD[12], Phil Hirsch, PhD[13], Patricia J. Jordan, PhD[14], Thomas J. Kim, MD, MPH[15], David D. Luxton, PhD[2], Michael D. Lynch, PhD, ABPP[16], Marlene M. Maheu, PhD[17], Francis L. McVeigh, OD, MS[18], Eve-Lynn Nelson, PhD[19], Chuck Officer[20], Patrick T. O'Neil, MD[21], Lisa J. Roberts, PhD[22], Colleen Rye, PhD[23], Carolyn Turvey, PhD[24], Alexander Vo, PhD[25]

[1]University of Colorado Anschutz Medical Campus, Aurora, Colorado; [2]Telehealth Program, National Center for Telehealth and Technology, Joint Base Lewis-McChord, Washington; [3]American Telemedicine Association, Washington, DC; [4]Department of Family Medicine and Community Health, University of Cincinnati, Cincinnati, Ohio; [5]University of Tennessee Health Science Center, Memphis, Tennessee; [6]Department of Psychiatry, University of Ottawa, Ottawa, Ontario, Canada; [7]Department of Psychiatry, University of Louisville, Louisville, Kentucky; [8]School of Health and Rehabilitation Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania; [9]Defense Health Headquarters, Falls Church, Virginia; [10]United States Air Force MSA, Arlington, Virginia; [11]Department of Psychiatry, University of Arkansas for Medical Sciences, Little Rock; [12]Physiology Associates, LLC, Frederick, Maryland; [13]Access Psychiatry Solutions, LLC, Seattle, Washington; [14]Pacific Telehealth & Technology Hui, Honolulu, Hawai'i; [15]Johns Hopkins Hospital, Baltimore, Maryland; [16]KACC/NRMC, Arlington, Virginia; [17]TeleMental Health Institute, Inc., Cheyenne, WY; [18]Telemedicine and Advanced Technology Research Center, Ft. Detrick, Maryland; [19]Department(s) Pediatrics and Kansas Center for Telemedicine & Telehealth, University of Kansas Medical Center, Kansas City, Kansas; [20]UnitedHealth Group, Minnetonka, Minnesota; [21]Department of Psychiatry and Behavioral Sciences Tulane University School of Medicine, New Orleans, Louisiana; [22]Bayer HealthCare, Bellevue, Washington; [23]Office of the Army Surgeon General, Arlington, Virginia; [24]University of Iowa, Iowa City, Iowa; and [25]Colorado Access, Denver, Colorado

Abstract

Background: The purpose of this document is to provide initial recommendations to telemental health (TMH) professionals for the selection of assessment and outcome measures that best reflect the impacts of mental health treatments delivered via live interactive videoconferencing.

Materials and Methods: The guidance provided here was created through an expert consensus process and is in the form of a lexicon focused on identified key TMH outcomes.

Results: Each lexical item is elucidated by a definition, recommendations for assessment/measurement, and additional commentary on important considerations. The lexicon is not intended as a current literature review of the field, but rather as a resource to foster increased dialogue, critical analysis, and the development of the science of TMH assessment and evaluation. The intent of this lexicon is to better unify the TMH field by providing a resource to researchers, program managers, funders, regulators and others for assessing outcomes.

Conclusion: This document provides overall context for the key aspects of the lexicon.

Keywords: telemental health, telepsychiatry, telemedicine, telehealth

Background

The realization of the need in telemental health (TMH) for a research lexicon to aid TMH professionals in the selection of assessment and outcome measures has evolved over the past several years within the membership of the American Telemedicine Association's (ATA) TMH Special Interest Group (SIG) and associated professional networks. This has been fostered by a series of parallel developments. First, the substantial increase in the utilization of videoconferencing during the past two decades promoted maturation in methodologies to describe, assess, and measure its impacts. As in any specialized field of medicine, the anticipated methodological evolution of studies began with pilot and descriptive studies, feasibility trials, and subsequently progressed to more rigorous assessments and outcome investigations such as randomized controlled trials.[1] Second, the current rapidly changing healthcare environment (with inherent uncertainties) imparts added urgency to these developments to include the need to better define the value of TMH interventions when communicating with funding organizations and other decision-makers in healthcare. This requires increased documentation of the quality, costs, and offsets of this modality, as well as equivalency in quality and outcomes compared to in-person care. It is these factors that have brought scrutiny and attention to the strengths, weaknesses, and gaps of the current evidence base that supports the use of TMH.

The field of TMH does have a burgeoning literature base, however, there are generally acknowledged limitations to the research.[2] Some of the more notable limitations are a modest number of outcome studies and a paucity of accompanying economic analyses. Many studies of TMH are conducted as pilot programs with small numbers of participants; and there are a limited number of rigorous comparative studies designed with adequate power. Additionally, while supporting the widespread use of TMH, the diversity of populations, age groups and settings represented in the published literature, creates problems for generalization of outcomes.[3] This is further complicated by the diverse array of study designs, methodologies, and measures employed to examine TMH programs and services. For example, while there are well over 70 published articles commenting on or examining patient/provider satisfaction in TMH, they employ different measures and models of satisfaction. This diversity of studies in the literature poses a challenge to building a broad set of data to aid in the integration of TMH into the larger healthcare system. This has propelled a growing consensus to build an evidence-base founded on clear and common definitions of assessment and outcomes that speaks to the specific impacts of TMH services. A common lexicon in this area will help grow understanding in this area, allow for broader comparisons, and support better generalization of findings.

Consensus Process

The process to develop an expert consensus lexicon used a previously demonstrated method for collaboration and consensus building developed by the University of Nebraska's Center for Collaboration Science. An overview of the workshop process follows, with a more detailed description provided by Mishkind, et al.[4] This workshop process used an electronic meeting system to more efficiently collect ideas and opinions, review comments, and craft a consensus document. The initial conceptualization was to constitute a one day workshop in conjunction with the ATA's 2012 Fall Forum meeting. Although significant progress on this document was made at the workshop, time limitations inhibited the production of a fully completed and finalized document, and the process was further modified to allow for completion of this work. The consensus workshop was designed by a small working committee constituted in the winter of 2012 and consisting of ATA TMH SIG Leaders, ATA Staff, and the director from the University of Nebraska's Center for Collaborative Science. In a series of teleconferences, draft documents and e-mails, this group developed a workshop agenda which included a packet of briefing materials, an invitee list, an outline of key areas/topics to be addressed by the lexicon and a process outline for the workshop. The invitee list was created by the group drawing from known experts from the field both within and outside the ATA membership, with a target goal of 25-35 participants. Invitations sent to potential participants provided an overview of the goals, expectations, and processes of the workshop. For those who accepted, a "briefing" packet was sent via e-mail 3-4 weeks prior to the workshop and consisted of key articles and an initial topic/area outline. Attendees were asked to review the packet and come to the workshop prepared to comment on the initial outline of topic/focus areas. Attendees represent a wide range of TMH expertise and backgrounds including representation from researchers, clinicians, policy experts, educators, academia, military, and private payers with experience in a range of populations (children, adults) and settings.

The full day workshop occurred in New Orleans on Sunday, September 9, 2012 a day preceding the ATA Fall Forum. Twenty six subject matter experts and two observers represented a wide range of backgrounds and depth of experience in TMH, including experts in administration, funding, research, program management and policy. The Director of the University of Nebraska's Center for Collaborative Science and a colleague conducted the workshop.
The day-long meeting opened with group introductions, a restating of the purpose, and an open discussion that focused on reaffirming consensus around the workshop goals. The workshop then facilitated a successive series of individual and small group brainstorming sessions to develop and refine the key items and their components. These were then reviewed by the entire group to provide specific commentary, edits, and opinions in an anonymous fashion using the software program. At the end of each of these developmental, brainstorming sessions the group reconstituted to discuss controversial items, review current status and reflect upon the group process. This process allowed for public discourse about the items. To achieve the completed lexicon, the workshop went through this iterative process four times, followed by anonymous voting on agreement and ranking of individual items. The entire process was supported by an electronic meeting system in which all participants were issued a computer loaded with the consensus process software.

It was determined that a definition, recommendations for assessment/measurement, and additional commentary on important considerations would be created for each identified item. During the first and second cycles, the group reviewed the key items/areas and then reached consensus on the final version of the components created for each item. In the third cycle, small groups of 2-3 participants, self-selected to work together, created the initial components for up to 3 items per group, followed by a review by all participants. The fourth cycle focused on revising and further crafting the items and their components based on the feedback from the third cycle. The workshop concluded with voting on level of agreement by item, ranking individual items, and discussion of next steps. Additionally, near the conclusion of the workshop it was realized that patient safety had not been addressed and a small group was constituted to produce this item.

As the workshop progressed throughout the day, the group came to the conclusion that the original goal of having a fully completed lexicon by the end of the workshop was overly ambitious. Although participants had high concurrence on the overall content of the lexicon, all agreed that it needed further refining of presentation, wording, and format. Post-workshop next steps were discussed to include having a primary writing team take the raw material from the workshop and further develop the document, which would then be sent out to all workshop participants for individual review and comment. The writing team would then incorporate these comments into a final draft.

Results

The major result of this effort is the lexicon of *assessment and outcome measures* presented in Table 1. Each item consists of (1) item definition; (2) item measurement recommendations; and (3) item measurement considerations.

Participants also used the consensus software to vote anonymously on whether they agreed with the components of each lexicon item by providing a yes/no vote to the question “Is this revised draft acceptable?” Results show an extremely high level of consensus across all items (see Table 2). Out of the 36 total items 13 had perfect agreement with all 26 participants concurring (100%), 9 items with 25 participants (96%), 10 items with 24 participants (92%), and 4 items with 23 participants all concurring (88%). The overall agreement rate across all items and participants was 96%.

As a final exercise, participants were asked to select what they felt were the key or essential outcome items (Table 3). The top ranked six items, each selected by over 65% of participants, were: quality of care (87%); patient satisfaction (78%); cost avoidance (70%); provider satisfaction (65%); usability (65%); and symptom outcome (65%). Four items were endorsed by less than 10% of participants as key outcomes: number of services (4%); public vs. private (4%). Neither supplies nor length of session were selected as essential by any participant.

Of note, item 2.1.3 (Tables 2 & 3), improved coordination or integration of care, was split into two items in the lexicon for clarification and simplification. The group decided item 2.3.7 (Tables 2 & 3) digital access, was sufficiently covered in item 2.3.6, cultural access and did not need to be included in the final lexicon.

Discussion

In addition to creating a needed tool for the field of TMH, the process of developing the lexicon provided a number of important lessons about guideline development. Most significantly, it demonstrated the utility and feasibility of consensus process which was adapted and employed for subsequent ATA practice guidelines development. Project leaders perceived the process to be superior to the previous approach wherein smaller committees worked via conference calls and drafting through asynchronous means. The consensus process appeared to be able to involve more people, occur more rapidly, and develop a consensus document driven by the group rather than 1-2 key leaders. The process allowed rapid and democratic decisions to be made by the entire group regarding 1) the key domains to be assessed and 2) the relative priority of those domains in demonstrating the value and impact of TMH. This particular experience also demonstrated the need to clearly develop realistic goals and expectations for the workshop. For example, expanding workshop time from 1 to 2 days with more predevelopment efforts (i.e., asynchronous outline drafting and discussion with the full workgroup regarding the workshop and expectations) might have initially produced a more refined lexicon. ATA took lessons learned from this experience and adapted the process for developing practice guidelines for telemedicine/telehealth. Figure 1 illustrates the revised process.

Several major themes appeared to emerge across items in the lexicon. Although arguably clinical treatment rendered via videoconferencing is often not substantively different than in-person care[3,5,6] there are unique aspects to TMH services that are not captured in traditional appraisal methods such as training, technology, and travel costs related to videoconferencing services. Thus, the importance of clearly defined measures and metrics along with the advice to use available validated measures speaks to the core goal of the lexicon. Thoughtful and appropriate selection of comparison groups is a theme that is also reiterated throughout the literature and not surprising to find in the lexicon.[3,5,6] Finally, it is striking that so many of the recommended measures are intended to examine the interaction of TMH service to the larger systems of care in which they are embedded.

Although this document and the workshop were unable to more fully address the topic of methodological models of research, the initial item grouping exercise on key outcomes for TMH (Table 3) provides some interesting guidance in this area. The top six ranked items by participants address central areas pertinent to healthcare reform and to the sustained implementation of TMH programs.[1, 4] The bottom four items provide important information at the individual program and project level - (1) number of services, (2) public vs. private, (3) supplies, and (4) length of session - but are measurements that carry less weight when assessing the ultimate, and broad-reaching, value of a service.

The concept of using a lexicon to bring unity to a field has precedent and is timely given the recent, April 2013 publication of *Lexicon for Behavioral Health and Primary Care Integration* by Agency for Health Research and Quality (AHRQ).[6] This lexicon much broader in scope, developed across two groups over a 4-year period, utilizes considerably more resources and time through the more traditional expert consensus approach. This speaks to both the consensus processes employed for the TMH Lexicon as well as supporting the general lexicon concept.[7]

This lexicon represents an important step in the continuing evolution in the field of TMH. Its impact will be determined over time by its reception in the larger TMH community. For progress to occur, individual researchers, assessors, and evaluators will need to use, test, validate, refine, and adapt the lexicon in their work. Such an evolution is crucial if TMH is to continue to improve the quality and access to care for patients in the context of the current healthcare environment.

References

1. America IoMCoQoHCi. Crossing the quality chasm: A new health system for the 21st century: National Academies Press; 2001.

2. Kramer GM, Shore JH, Mishkind MC, Friedl KE, Poropatich RK, Gahm GA. A standard telemental health evaluation model: the time is now. *Telemed J E Health*. 2012; 18(4):309-13.

3. Grady B, Myers KM, Nelson EL, Belz N, Bennett L, Carnahan L, et al. Evidence-based practice for telemental health. *Telemed J E Health*. 2011; 17(2):131-48.

4. Mishkind MC, Doarn RC, Bernard J, Shore J. The Use of Collaboration Science to Define Consensus Outcome Measures: A Telemental Health Case Study. *Telemed J E Health* 2013; 19(6):455-59. Epub Apr 16

5. Myers K. Turvey C. Editors. Telemental Health: Clinical, Technical and Administrative Foundations for Evidence-Based Practice. New York: Elsevier; 2013.

6. Yellowlees P, Shore J, Roberts L. Practice guidelines for videoconferencing-based telemental health - October 2009. *Telemed J E Health*. 2010;16(10):1074-89.

7. Council PCatNIA. Lexicon for Behavioral Health and Primary Care Integration: Concepts and Definitions Developed by Expert Consensus. Rockville, MD: Agency for Healthcare Research and Quality., 2013 Contract No.: AHRQ Publication No.13-IP001-EF.

Table 1. Lexicon of Assessment and Outcome Measures for Telemental Health (TMH)

Item No.	Item	Definition	Measurement	Considerations
2.1.1	Patient satisfaction	Patient's subjective satisfaction and experience with the TMH service provided.	The perception of the patient's satisfaction during the TMH visit with usability of the technology, patient-provider communication, and convenience of receiving care via this approach. Does the patient believe that the service met her/his health needs? Would patient do this again? Would patient refer others to this service?	There may be overlap with other constructs such as “Satisfaction with Usability of Technology." Satisfaction does not necessarily require in-person comparison. It could be comparison to no care (i.e., non-inferiority testing). Use of validated measures of TMH satisfaction because measures exist. Measure satisfaction with experience as well as with technology.
2.1.2	Provider Satisfaction	The extent to which the provider values telehealth when interacting with patients.	The following metrics may serve as surrogate markers: retention and recruitment of providers, ease of transition in technical competency, ease of integration into clinical workflow, perceived value of better diagnosis, treatment and disease management.	Satisfaction metric must be considered longitudinally. Include both referring PCMs and consulting provider satisfaction surveys.
2.1.3	Coordination of care	Care coordination is the development and implementation of a shared plan to support patient wellness.	Care coordination measurement consists of both the number of telehealth encounters and the number of different participants involved in the shared plan (e.g., consultant-primary care provider, consultant-teacher, etc.) and the type of telehealth interaction (asynchronous and synchronous).	The nature of the communication, external technologies such as electronic health records and quality of encounters can all impact care coordination.
2.1.3	Integration of care	Integration of care is the efficient assimilation of	Measurement of the integration of care includes the type of the	The nature of the communication, external technologies such as electronic health records and

		multiple components within a health system in order to decrease redundancy, delay, and cost.	telehealth interactions assessed on standardized questionnaires of care coordination or other measures of communication (i.e., participant A to participant B).	quality of encounters can all impact integration of care.
2.1.4	Usability	1) The ease (preference, comfort, fit, readiness) of patients to communicate digitally with their providers. 2) Includes technology availability, simplicity of use, service availability, technology native vs. non-facile.	Measurement should include: provider retention rate, patient drop out and rationale, support staff required, technology ease of use, technology down time, and subjective ratings of comfort.	Subjective and objective measurements from both the patient and provider perspective. Part of the evaluation should include how "seamless" the interaction was between people/technology, to include latency and failure of technology. This can be used as both a process/acceptability and an access measure, but definition should remain the same. Patient/provider preferences should also be included.
2.1.5	Rapport	When two or more people feel that they are connected and understand one another.	Self-reported level of direct and/or indirect evidence that the condition of rapport is present between the patient(s) and the professional(s).	Transcends cultural, racial, ethnic, religious, gender, age, geographic, etc. differences and experiences. Try to link clinical outcomes which could be related to rapport.
2.1.6	Stigma	Preconceived, often negative, association with an illness, diagnosis, therapy, technique etc. that may interfere with the provision and/or acceptance of care.	Measures should evaluate stigma among health care providers/staff, patients, and social networks and include, at minimum, the following concepts: Stereotyping/discrimination such as beliefs about mental illness, mental health treatment, TMH and the use of technology to deliver care. Labeling/disclosure such as acceptance of diagnosis, willingness to diagnose appropriately, help seeking and delivering behaviors, willingness to use or conduct TMH	Perceived stigma should not simply focus on the recipient of care but the providers of care and those giving support. Concerns about stigma should focus on both mental illnesses in general and on the type of delivery (e.g., TMH). From a research and programmatic perspective this is best evaluated pre/post introduction of a TMH service. This can be related to both general access to care and readiness.

			sessions.	
2.1.7	Motivational readiness	Assessment of an individual's or organization's willingness to change and adopt TMH services. This is different from preparedness, which is an assessment of individual and organizational ability to adopt TMH services.	Includes: stage of change for individuals and organizations, situational self-efficacy (confidence), trans theoretical model-based measures (pros & cons of change, processes/strategies for change, situational self-efficacy).	Defining criteria for moving into the action stage. Relationship between individuals and institutional readiness and motivation. How interrelated are individuals and institutional motivation? Self-report can be inaccurate, but necessary.
2.2.1	No shows	A patient or clinician who does not attend session, or is more than 15 minutes late.	Percent of no shows as compared to a disease-state specific comparisons in-person group. No shows defined as 15 minutes late or more to appointment. No shows need to be identified as either clinically related or a systems issue (scheduling, time zones, etc.).	Determine cause of no show, i.e., was it lack of transportation, lack of ability to maintain a schedule, did they show up late and have to reschedule, dissatisfaction with treatment. Examine the reasons for the no shows i.e. technology failed or could not be used, the use of technology (vs. travel) made it easier to keep the appointment, etc.
2.2.2	Accuracy of assessment	How well the modality of TMH impacts the reliability and validity of the assessment when compared with the traditional behavioral health care standards for the construct in question.	Comparison of standard measures of assessment (reliability, validity) of TMH vs. in-person (national standard) vs. other telehealth modalities. Measurement should also include session time and number of sessions needed for specific assessments comparing TMH with in-person services at patient site.	Proxy measures to track providers comfort with reliability of assessment through tracking utilization of tests and consults comparing TMH with in-person services at patient site.
2.2.3	Symptom outcomes	Change in identified clinical symptoms over time.	Use of measures of symptom change that are appropriate and psychometrically sound (validity, reliability data published in the literature). Need to be appropriate	How is this information documented so it is meaningful? Include measure used, cutoff criteria, inclusion/exclusion, what they are comparing outcome to, effect size of intervention. Symptom outcomes are part of a larger universe of outcome

			for the population being treated/assessed to include accepted gold standards.	metrics that need to be considered. Consider adding intervention/treatment outcomes with symptom outcomes as a subset as well as other outcomes such as Quality of Life, work attendance/absenteeism, compliance/adherence or psychosocial measures (unit cohesiveness, social isolation).
2.2.4	Completion of Treatment	Degree to which appointments, treatments and completion of treatment plans occurred within the prescribed time frame.	Average number of visits according to treatment plan, average number of visits in given time period, duration of treatment, number/percentage of modules completed; percentage of patients who completed treatment; pre/post functional measures	Third party payers use Axis 5 (Global Assessment of Functioning) to evaluate progress and completion, although this will evolve with the conversion to DSM-V criteria.
2.2.5	Quality of Care	Quality of care represents the process of delivering services and includes both the technical and interpersonal aspects of treatment. Technical quality includes concordance with treatment guidelines, fidelity to evidence based protocols, and system performance measures (e.g., HEDIS). Interpersonal quality includes patient rapport, therapeutic alliance, and cultural competence.	Performance measures (e.g., timely outpatient visit follow hospital discharge) can sometimes be measured from administrative data. Concordance with treatment guidelines and fidelity to evidence based protocols can be measured from chart review. Interpersonal quality should be measured from patient self-report (e.g., therapeutic alliance can be measured using the working alliance inventory).	Quality is defined as the process rather than the outcome of care, because clinical outcomes are measured using other metrics and because high quality care does not necessarily lead to good outcomes. Quality of TMH services should be measured against benchmarks rather than the quality of in-person services which is often sub-optimal. When TMH services are compared to in-person services, it will be critical to choose a similar clinical setting and patient population.
2.2.6	Treatment	Use of TMH services	Measurements on number of TMH	Comparison of digital contacts (mobile phone, e-

	Utilization	compared with all other health services related to specific disease processes.	and non-TMH visits within a health care system to include data on visit duration, frequency, and problem addressed. Measurements on system resources (labs, medications, system funded travel, devices, consultation, number of referrals made and utilized) of TMH vs. non-TMH. Utilization should be correlated with symptom reduction of specific disease processes.	mail, Web) and its impact on service utilization in non-telemental healthcare. Recommend healthcare systems systematize data on digital contacts. Collect data on both internal utilizations within a system but as possible external service utilizations from outside agencies and providers. As possible during implementation of TMH services collect compare data on pre and post implementation service utilization data.
2.3.1	Number of Services	Degree of access to additional services which are derived from enrollment in telehealth.	The number of clinical care options and auxiliary services offered (e.g., medication management, social services, labs, cardiac care, group therapy); frequency in the use of clinical care options and auxiliary services.	Used for program evaluation, ROI for program expansion, quality, patient/provider satisfaction.
2.3.2	Numbers Served (also referred to as RVUs, relative value units)	The workload credit given for the TMH encounter that is related to the complexity of services provided and the time spent with patients which equates to the level of financial reimbursement.	Types of services; complexity of services; time spent with patients; number of patients seen.	Coding accuracy. Coding training and follow up to ensure coding is being done correctly, i.e., no under or over coding.
2.3.3	Wait Times	Wait time is a temporal dimension of access that represents the delay between when the patient wants to receive	Operationally, time to next available appointment, when scheduling, and when the patient actually presents for care. For TMH requiring a referral, wait time could be measured as the	It is important to realize that improving other dimensions of access (e.g., lowering costs or de-stigmatizing TMH services) could result in increasing wait times due to increased demand. Health systems should measure wait times to all

		services and when they can actually receive services.	difference in the referral date and the date the patient was seen. May want to measure wait time separately to see the preferred provider versus any provider.	clinics (not just TMH clinics) to determine how resources could best be reallocated to minimize variability in wait times across clinics. Other important measures of temporal access include wait time in clinic and convenience of office hours.
2.3.4	Length of session	How much time the patient spends receiving care. This could include time spent with the provider.	Average/total clinical encounter time, average/total administrative time (set-up time, out-of session contact such as email, text, phone, letters).	Needs to be clinician, patient, staff, and system viewpoint. Needs to accommodate emerging platforms such as mobile health. Length of sessions may interact with frequency of appointments. Efficiencies with telehealth solution create opportunities for novel session duration (e.g., 10-minute check-in)
2.3.5	Distance to Service	Geographic separation or functional barriers between patients and providers.	Distance, time zones, time to appointment.	This includes structural barriers, weather.
2.3.6	Likelihood to access vs. traditional care	Likelihood to use TMH.	Measurement should include the following concepts: familiarity (past use), acceptability (cultural and technical), associations with stigma, willingness, and perceived benefit. Measurement should not focus on satisfaction but rather broad willingness to use.	When possible this should include baseline comparisons against both available and unavailable treatment as usual (e.g., in-person) Most likely this is assessed through self-report questionnaires.
2.3.8	Cultural access	Access to healthcare services that align with cultural expectations.	The degree to which an individual perceives the mode of delivery and related processes to align with cultural beliefs and expectations.	This should include cultural understanding of technology and expectations of interpersonal communication. It should also consider how technology may better connect cultural expectations, e.g., providing access to same culture providers or allowing for communication with a provider outside of one's in-group.
2.4.1	Economic evaluation that incorporates			In general, clear definitions do not exist for many of the cost structures. This may be appropriate as costs are derived and perceived differently. There

	standard economic models			are several costs factors that were identified as important to measure objectively. Until final definitions are set, each cost factor should be operationalized and reported. Consideration should also be given to what is sunk or similar cost of care as usual (provider time). Baseline assessments help to identify cost outcomes.
2.4.2	Value proposition	Comparison of clinical and other health service outcomes by overall resources allocated.	Standardized and reported taxonomy of resources allocated and outcomes measured.	There is no consensus yet on the best determinations for economic evaluations in TMH.
2.4.3	Travel direct	Direct cost associated with provider and/or patient travel to care site	All direct costs should be identified, operationalized, and reported for comparison.	Should be included within the broad category of costs. Precise definition may not be possible given differing perspectives but all components should be identified, operationalized, and reported.
2.4.4	Travel indirect	Indirect costs associated with provider and/or patient travel to care site	All indirect costs should be identified, operationalized, and reported for comparison.	Should be conceptualized as comparison to normal care, e.g., loss of work productivity is comparable given 1 hr away regardless of mode of delivery. Indirect costs are both inputs to a cost model as well as potential positive outcomes of telehealth (reduction). Evaluators should determine and report up-front whether indirect costs are inputs to a cost model or expected outcomes.
2.4.5	Technology direct	Direct patient and provider costs associated with the technology utilized to deliver telehealth services.	All direct costs should be identified, operationalized, and reported for comparison.	Need to determine upfront whether costs are as a whole or divided between provider- and patient-associated. Inputs to consider include: hardware and depreciation, software and licensing, infrastructure, network, and maintenance costs.
2.4.6	Technology indirect	Indirect patient and provider costs associated with the	Indirect costs include expenses incurred as a result of technology downtimes, specialized licenses, and	There is cross-over between direct and indirect technology costs. Direct costs should focus on tangible assets while indirect costs are often

		technology utilized to deliver telehealth services.	administration.	intangible resources allocated based on the need for tangible assets.
2.4.7	Public vs. private	Payer Perspective.	Whether a project, program, or system utilizes public or private funding.	This is not an outcome measure but rather a perspective. Outcomes measures should be evaluated based upon the financial perspective under which a program operates.
2.4.8	Cost avoidance	Current or future direct costs avoided due to a specific intervention or program.	There are currently no industry standards for cost avoidance measures.	Consideration should be given to measuring items such as hospitalizations, visits, and other costs. These should be operationalized and reported as possible.
2.4.9	Missed obligations	Indirect Cost: Missed obligations	Should be measured as part of overall indirect costs.	Where possible a baseline assessment should be conducted against care as usual. As an outcome measure the assumption is that TMH impacts indirect costs/burden, thus requiring a comparison.
2.4.10	Burden on social network	Societal resources associated with either the provision of or inadequate access to TMH services.		Burden on social network should include direct burden to support resources and broad burden to societal infrastructure. When conducting research a positive or negative directional association should be identified *a priori.*
2.4.11	Personnel (administrative, provider, provider extender, presenter)	Personnel costs associated with the provision of TMH services.		
2.4.12	Supplies	Direct cost of auxiliary supplies required for TMH services.		
2.4.13	Training	Process by which an individual attains the knowledge and skills required to	A TMH competency set is required.	May be included as an indirect provider cost. Training is not truly an outcome unless the program is development of a training program

		demonstrate predetermined competencies.		
2.4.14	Facilities and maintenance	Direct costs associated with the facilities and maintenance necessary to support telehealth technologies.	Measurement includes cost of physical facilities, facilities maintenance, and systems such as HVAC. Should also include cost to maintain equipment including servers and individual patient/provider technologies.	Should be included with technology direct costs.
2.4.15	Broad resource utilization	Resource utilization is the total allotment of resources necessary to provide telehealth services.	Resource utilization is driven by the numbers of encounters. It encompasses personnel and infrastructure resources necessary to provide each health care service.	Baseline comparisons need to be considered to differentiate resources from treatment as usual and TMH.
Patient safety	Patient safety	Safety of patients and others during the course of treatment (i.e. during sessions and after).	Times had to use safety procedures. Number of times needing to contact collateral/ 911 calls/emergency services calls. Number of psychiatric hospitalizations related to clinic services. Number of times unable to invoke safety plan (tried but could not), hand off to higher level of care from clinic due to safety issues. Problems causing patient transfer to another provider.	Consider Targsoff, other measures of adverse events (or potential ones e.g. increased suicide indication, etc.); response times of all events, etc. including emergency services.

Table 2. Voting results on question: Is this revised draft acceptable? (Y/N)

		Vote Distribution					
#	Ballot Items	N	Y	Avg. Score	Total	STD	Votes
1	2.1.1. Patient satisfaction	-	26	1.00	26.00	0.00	26
2	2.1.2. Provider satisfaction	-	26	1.00	26.00	0.00	26
3	2.1.3. Improved coordination or integration of care	1	25	0.96	25.00	0.20	26
4	2.1.4. Usability	3	23	0.88	23.00	0.33	26
5	2.1.5. Rapport (comfort, cultural competence, language barriers)	2	24	0.92	24.00	0.27	26
6	2.1.6. Stigma	-	26	1.00	26.00	0.00	26
7	2.1.7. Motivational readiness	2	23	0.92	23.00	0.28	25
8	2.2.1. No shows	-	26	1.00	26.00	0.00	26
9	2.2.2. Accuracy of assessment	-	26	1.00	26.00	0.00	26
10	2.2.3. Symptom outcomes	-	26	1.00	26.00	0.00	26
11	2.2.4. Completion of treatment	-	26	1.00	26.00	0.00	26
12	2.2.5. Quality of care	-	25	1.00	25.00	0.00	25
13	2.2.6. Treatment utilization	2	24	0.92	24.00	0.27	26
14	2.3.1. Number of services	2	24	0.92	24.00	0.27	26
15	2.3.2. Numbers served	3	23	0.88	23.00	0.33	26
16	2.3.3. Wait times	1	25	0.96	25.00	0.20	26
17	2.3.4. Length of session	1	25	0.96	25.00	0.20	26
18	2.3.5. Distance to service	2	24	0.92	24.00	0.27	26
19	2.3.6. Likelihood to access vs. traditional care	3	23	0.88	23.00	0.33	26
20	2.3.7. Digital access	1	24	0.96	24.00	0.20	25
21	2.3.8. Cultural access	2	24	0.92	24.00	0.27	26
22	2.4.1. Economic evaluation that incorporates standard economic models	1	24	0.96	24.00	0.20	25
23	2.4.2. Value proposition	1	24	0.96	24.00	0.20	25
24	2.4.3. Travel direct	-	26	1.00	26.00	0.00	26
25	2.4.4. Travel indirect	-	26	1.00	26.00	0.00	26
26	2.4.5. Technology direct	-	26	1.00	26.00	0.00	26
27	2.4.6. Technology indirect	1	25	0.96	25.00	0.20	26

28	2.4.7. Public vs private	2	24	0.92	24.00	0.27	26
29	2.4.8. Cost avoidance	-	26	1.00	26.00	0.00	26
30	2.4.9. Missed obligations	1	25	0.96	25.00	0.20	26
31	2.4.10. Burden on social network	1	25	0.96	25.00	0.20	26
32	2.4.11. Personnel (administrative, provider, provider extended, presenter)	-	26	1.00	26.00	0.00	26
33	2.4.12. Supplies	2	24	0.92	24.00	0.27	26
34	2.4.13. Training	1	25	0.96	25.00	0.20	26
35	2.4.14. Facilities and maintenance	-	26	1.00	26.00	0.00	26
36	2.4.15. Broad Resource Utilization	-	25	1.00	25.00	0.00	25

Table 3. Number of times an outcome was selected to be part of the 'key group of outcomes' (0 = not selected, 1 = selected)

		Distribution					
#	**Ballot Items**	**0**	**1**	**Avg. Score**	**Total**	**STD**	**Votes**
1	2.1.1. Patient satisfaction	5	18	0.78	18.00	0.42	23
2	2.1.2. Provider satisfaction	8	15	0.65	15.00	0.49	23
3	2.1.3. Improved coordination or integration of care	12	11	0.48	11.00	0.51	23
4	2.1.4. Usability	8	15	0.65	15.00	0.49	23
5	2.1.5. Rapport (comfort, cultural competence, language barriers)	16	7	0.30	7.00	0.47	23
6	2.1.6. Stigma	16	7	0.30	7.00	0.47	23
7	2.1.7. Motivational readiness	19	4	0.17	4.00	0.39	23
8	2.2.1. No shows	19	4	0.17	4.00	0.39	23
9	2.2.2. Accuracy of assessment	10	13	0.57	13.00	0.51	23
10	2.2.3. Symptom outcomes	8	15	0.65	15.00	0.49	23
11	2.2.4. Completion of treatment	14	9	0.39	9.00	0.50	23
12	2.2.5. Quality of care	3	20	0.87	20.00	0.34	23
13	2.2.6. Treatment utilization	13	10	0.43	10.00	0.51	23
14	2.3.1. Number of services	22	1	0.04	1.00	0.21	23
15	2.3.2. Numbers served	14	9	0.39	9.00	0.50	23
16	2.3.3. Wait times	15	8	0.35	8.00	0.49	23
17	2.3.4. Length of session	23	-	0.00	0.00	0.00	23
18ss	2.3.5. Distance to service	12	11	0.48	11.00	0.51	23

19	2.3.6. Likelihood to access vs. traditional care	17	6	0.26	6.00	0.45	23
20	2.3.7. Digital access	21	2	0.09	2.00	0.29	23
21	2.3.8. Cultural access	15	8	0.35	8.00	0.49	23
22	2.4.1. Economic evaluation that incorporates standard economic models	12	11	0.48	11.00	0.51	23
23	2.4.2. Value proposition	15	8	0.35	8.00	0.49	23
24	2.4.3. Travel direct	15	8	0.35	8.00	0.49	23
25	2.4.4. Travel indirect	18	5	0.22	5.00	0.42	23
26	2.4.5. Technology direct	17	6	0.26	6.00	0.45	23
27	2.4.6. Technology indirect	19	4	0.17	4.00	0.39	23
28	2.4.7. Public vs private	22	1	0.04	1.00	0.21	23
29	2.4.8. Cost avoidance	7	16	0.70	16.00	0.47	23
30	2.4.9. Missed obligations such as work	20	3	0.13	3.00	0.34	23
31	2.4.10. Burden on social network	20	3	0.13	3.00	0.34	23
32	2.4.11. Personnel (administrative, provider,provider extended, presenter)	18	5	0.22	5.00	0.42	23
33	2.4.12. Supplies	23	-	0.00	0.00	0.00	23
34	2.4.13. Training	16	7	0.30	7.00	0.47	23
35	2.4.14. Facilities and maintenance	22	1	0.04	1.00	0.21	23
36	2.4.15. Broad resource utilization	20	3	0.13	3.00	0.34	23

Figure 1. ATA Practice Guidelines Consensus Workshop Process