



## Evidence-based Practice Center Technical Brief Protocol

### Project Title: *Evaluation of Mental Health Applications*

#### I. Background and Objectives for the Technical Brief

Among adults aged 18 or older in the U.S., the prevalence of mental illnesses has increased from 17.7% (39.8mil Americans) in 2008 to 20.6% (51.5mil) in 2019.<sup>1</sup> Of this population, 26.0% (13.3mil) perceived having an unmet need for mental health services. The most common reason for unmet needs is affordability, with other barriers including mental health stigma, and provider shortages or wait time.<sup>2</sup>

Evidence-based mobile health (mHealth) interventions have great potential to provide much needed access to care. MHealth interventions can be readily disseminated with minimal need for staff training, resource investment, and with limited burden on healthcare settings. Furthermore, technology-enhanced healthcare approaches are increasingly being reimbursed by insurers, making apps more likely to be integrated into the toolkits of healthcare providers. Apps could also help to address ethnic/racial and geographic disparities in access to and uptake of mental health services. Adjunctive treatment with apps is affordable, offers a way to reach rural populations and traditionally hard-to-reach groups, and can reduce patients' feelings of stigmatization or discomfort with traditional treatment.<sup>3</sup>

Several systematic reviews have identified positive effects of mental health apps on smoking behavior, self-reported depressive symptoms, and anxiety scores.<sup>4-6</sup> However, the effectiveness of most available mental health apps is not supported by research. This lack of evidence, combined with uncertainty surrounding safety and privacy, and the lack of insurance coverage or reimbursement for the majority of mental health apps, causes providers to be hesitant about recommending apps.<sup>7</sup> Most providers and patients need guidance in choosing the most appropriate app for their needs.

Growing concerns over data sharing<sup>8-10</sup> and clinical safety<sup>11</sup> have created a demand for an assessment framework assessing the clinical foundation, evidence-base, accessibility, purpose, security and privacy features, usability, and safety to help consumers, family members and peer supports, providers and health systems select mental health mobile apps. A study found that out of 36 top-ranked apps for smoking cessation and depression, 29 sent data to Facebook or Google, but only 12 apps mentioned this in their privacy policy.<sup>10</sup> Another study of 61 mental health apps found that 41% of the apps did not have a privacy policy at all.<sup>9</sup>

Given the uncertain evidence-base for most apps, safety/privacy concerns and other issues, the ***decisional dilemma is “how can consumers, family members and peer supports, providers and health systems select mental health and wellness mobile applications?”*** The aim of this project is to address this dilemma through two steps. The first step is to delineate the characteristics and minimum standards to inform the development of a framework to assess the appropriateness and effectiveness of mental health apps for users across different age groups and for different mental health disorders,

including general mental wellness. The second step is to evaluate the framework by applying it to a group of mental health and wellness apps.

## II. Guiding Questions

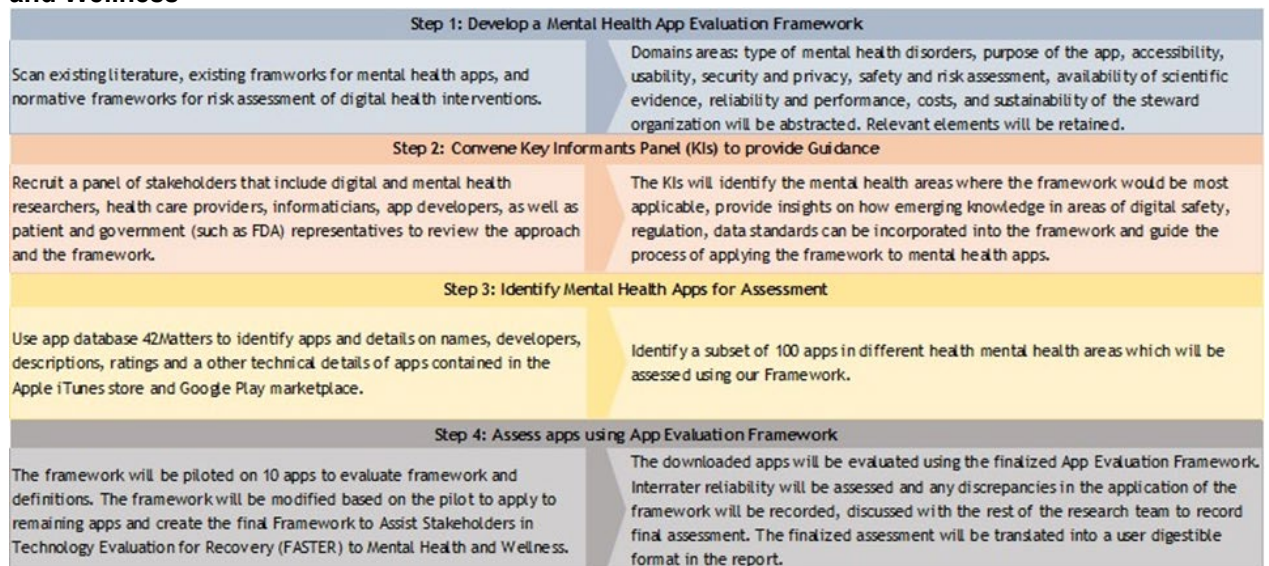
1. What characteristics and minimal standards of available behavioral health mobile applications need to be analyzed in existing tools to assess the appropriateness (to various stakeholders) and effectiveness of available apps to include, but not limited to:
  - Accessibility including ease of use, health literacy, 508 compliance, digital equity, cost
  - App background including funding source, purpose
  - Security features and privacy policy such as data ownership/usage
  - Clinical foundation and linkage to current evidence-base
  - Usability, including interoperability across platforms, stability
  - Therapeutic goals, linkage to the provider, crisis warning notification/alert system
  
2. Identify or develop an assessment framework for evaluation/scoring tools (e.g., websites) and apply the framework to help consumers, family members and peer supports, providers and health systems select behavioral health mobile applications. The framework will take into account current FDA status on the use and classification of risks of apps in healthcare.

Applications reviewed by the tools are intended to be used for screening, monitoring and management of mental health symptoms or disorders, response to treatment and to assist with general mental wellness. They can be intended for use by individuals of any age with any mental health condition (e.g., depression, anxiety, substance use disorders, post-traumatic stress disorder, bipolar disorder, psychosis, opioid use disorder), and mental wellness (e.g., mindfulness, meditation) for the general population.

## III. Methods

To address the aims of this Technical Brief, we will undertake a series of activities outlined in Figure below to develop the Framework to Assist Stakeholders in Technology Evaluation for Recovery (FASTER) to Mental Health and Wellness.

**Figure : Overview of the process to develop the Framework to Assist Stakeholders in Technology Evaluation for Recovery (FASTER) to Mental Health and Wellness**



## A. Discussions with Key Informants

We will recruit digital and mental health researchers, health care providers, informaticians and technologists, app developers, as well as patient and government (such as FDA) representatives to provide user and expert perspectives on how mental health apps work, how they might fit in clinical care, as well as potential advantages and disadvantages. We will conduct KI interviews via group calls on Zoom to elicit guidance on: (i) the scope of the project in terms of mental health conditions and wellness domains to which the framework can be applied; (ii) the ideal domains of the framework, from KI perspectives, including identifying gaps and providing recommendations to strengthen each domain of the framework; and (iii) the process of selection of the apps and application of the framework.

We will conduct KI calls with two different groups (group 1: clinical; group 2: tech, app developers). At the beginning of the project, the first group of KIs will be asked questions about apps and elements of potential framework, including conditions for which mental health apps are helpful, risks or concerns about use of apps, and key characteristics to consider in recommending/covering/choosing apps.. We plan to have a separate call for patient/caregiver KI at the start of the project, with questions modified from this first group.

The second KI group will be engaged 2-4 weeks after the meeting with the first KI groups, and after a draft framework has been developed. They will be sent a list of draft criteria/draft framework at least one week before the meeting and asked to provide insights on how emerging knowledge in areas of digital safety, regulation, evidence, data standards can be incorporated into the framework. For example, one key decisional criterion would be about leveraging existing

frameworks for evidence and risk assessment of digital health tools as part of this project. Prior to the pandemic, FDA regulated mental health apps through the Software Pre-Certification Pilot program that first evaluates the software developer or digital health technology developer, rather than the product. Other relevant frameworks include the UK's National Institute for Health and Care Excellence (NICE) Digital Health Technologies (DHT) Evidence Standards Framework<sup>12</sup> and the Department of Health & Social Care in UK "A Guide to good practice for digital and data-driven health outcomes" which has a section on generating evidence of clinical, social, economic or behavioral benefits.<sup>13</sup>

The second call will include technologists and app developers and focus on eliciting feedback on the developed framework. The questions will focus on aspects of the framework related to user interface/ease of use, functionality, and issues of security, privacy, risk mitigation and interoperability.

## **B. Gray literature search.**

We will scan grey literature based on the team's knowledge of existing frameworks such as One Mind PsyberGuide, VeryWell Mind, Health Navigator New Zealand, UK National Health Service, MyHealthApps, Health Living App Guide - Vic Health, and American Psychiatric Association App Advisor. We will review other normative frameworks for assessment such as UK's National Institute for Health and Care Excellence (NICE), and Digital Health Technologies (DHT) Evidence Standards Framework Policy for Device Software Functions and Mobile Medical Applications. We will leverage the frameworks that our team has already developed, specifically the most recently developed framework which is currently undergoing a consensus process with the WHO Digital Health Clearinghouse. We will also ask KIs for suggest resources. Supplemental Evidence And Data for Systematic review (SEADS) portal will be available for this technical brief.

## **C. Published literature search.**

We will conduct targeted searches of PubMed, and complete snowball searching, to identify existing frameworks for the evaluation of mental health apps. We will also ask KIs for suggested published literature. Our preliminary search strategy is in Appendix A.

## **D. Identification/development of framework.**

We will review each of the frameworks, identify common domains across the frameworks, and extract the relevant domains and their definitions. For each abstracted domain, the relevance to this project will be discussed and determined among the team. The rationale for the decision will be documented. The domains may include the following : type of mental health disorders and severity; purpose of the app; accessibility including health literacy, 508 compliance, digital equity, and cost; usability and acceptability by patients/clients and health care providers; security and privacy features such as data usage and sharing of data on social media as well as regulatory oversight; safety and risk assessment, availability of scientific evidence to support the intervention delivered by the app; reliability and

performance; and sustainability of the steward organization such as financial and resource stability, funding sources and future plans for growth and partnership with health delivery providers will be evaluated.

Assessment of risks posed due to the use of the app will be a critical element of the framework and will be based on severity of mental disorder of focus, the evidence of the clinical effectiveness of the app for the population of focus, and the ability of the app to connect to crisis services. From a technology and platform perspective, the framework will also evaluate interoperability with healthcare systems, use of standards, performance, and ability to scale to support large sets of users. Instructions for the application of the framework will be drafted.

## **E. Application of framework.**

We will use the app database 42Matters (<https://42matters.com/>) which provides a database with names, developers, descriptions, and a range of technical details of apps contained in the Apple iTunes store and Google Play marketplace. This database allows for the download of user feedback from app stores as well as app categories, pricing details, country of origin, ratings, downloads, release dates, and content localization. Informed by input from the KIs, we will use relevant search terms to search for mental health apps. Terms identified thus far include “depression”, “anxiety”, “alcohol abuse”, “binge drinking”, “smoking”, “insomnia”, “stress”, “Bipolar Disorder”, “eating disorder”, “mental illness”, “obsessive compulsive disorder”, “Post Traumatic Stress disorder”, “PTSD”, “psychotherapy”, “substance abuse disorders”, “psychosis”, “opioid use disorder”, “suicidal thoughts”, and “mental wellness” (e.g., mindfulness, meditation)”; we will search in the “Title” and “Description” categories of the applications. The search will be limited to Interactive Advertising Bureau (IAB) categories of “Medical Health” and “Healthy Living” from 2010 onwards that are in English and accessible in the U.S. The IAB develops industry standards to support standardization in the digital advertising industry and serve as a quality assurance guideline taxonomy to facilitate transparent marketing of publicly available apps.

National Institute of Mental Health (NIMH) classifies mental health apps into six categories based on functionality: self-management, cognition improvement, skills-training, social support, symptom tracking, and passive data collection. Mental health apps span all stages of clinical care provision, including immediate crisis intervention, prevention, diagnosis, primary treatment, supplement to in-person therapy, and post-treatment condition management. We will identify core mental health “buckets” within which apps will be selected for evaluation by the framework. These categories or “buckets” will be selected based on criteria which will be defined a priori, including: the mental health content area (e.g., anxiety, depression, addictions, insomnia); whether apps are available free of cost, for a subscription or with in-app purchases; the number of downloads; availability and quality of evidence on mental health and wellness outcomes; and adherence to FDA regulations. Based on the a priori criteria, apps will be screened for inclusion until the target number of apps (3-5) in each bucket is reached, not to exceed a total 100 apps.

The included apps will be downloaded. Team members will first conduct pilot testing of framework and instructions by applying the framework to 10-20 apps. This will be completed for the draft report. During peer review, the framework will be applied to the remaining apps. At least 2 members of the research team will apply the framework to 20% of the included apps. Depending on the structure of the final framework, apps will receive a score of “1” to indicate the presence of a certain feature, or a “0” to indicate the absence of the feature. Interrater reliability between the two reviewers will be assessed and any discrepancies in the application of the framework will be recorded, discussed with the rest of the research team, and definitions will be further clarified. We will then continue applying the framework to the rest of the apps, with one person applying the framework to each app. We will discuss the process of applying the framework to generate further refinements of the framework and instructions.

## **2. Data Organization and Presentation:**

### **A. Information Management**

We will use Excel to track relevant grey and published literature about frameworks, to list and describe characteristics of apps (guiding question 1) and to define and compare framework domains identified.

### **B. Data Presentation**

Tables will be presented of key characteristics of apps (guiding question 1) and the identified/developed framework will be presented (guiding question 2). Instructions regarding use of the framework will be provided, such as definitions and details on how to apply the framework.

## **IV. References**

1. Substance Abuse and Mental Health Services Administration. (2020). Key substance use and mental health indicators in the United States: Results from the 2019 National Survey on Drug Use and Health (HHS Publication No. PEP20-07-01-001, NSDUH Series H-55). Rockville, MD: Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration. Retrieved from <https://www.samhsa.gov/data/>.
2. Carbonell Á, Navarro-Pérez JJ, Mestre MV. Challenges and barriers in mental healthcare systems and their impact on the family: A systematic integrative review. *Health & social care in the community*. 2020 Sep;28(5):1366-79. doi: 10.1111/hsc.12968. PMID: 32115797.
3. Wang K, Varma DS, Proserpi M. A systematic review of the effectiveness of mobile apps for monitoring and management of mental health symptoms or disorders. *Journal of psychiatric research*. 2018 Dec;107:73-8. doi: 10.1016/j.jpsychires.2018.10.006. PMID: 30347316.
4. Huckvale K, Nicholas J, Torous J, et al. Smartphone apps for the treatment of mental health conditions: status and considerations. *Current opinion in psychology*. 2020 Dec;36:65-70. doi: 10.1016/j.copsyc.2020.04.008. PMID: 32553848.

5. Weisel KK, Fuhrmann LM, Berking M, et al. Standalone smartphone apps for mental health-a systematic review and meta-analysis. NPJ digital medicine. 2019;2:118. doi: 10.1038/s41746-019-0188-8. PMID: 31815193.
6. Firth J, Torous J, Nicholas J, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. World psychiatry : official journal of the World Psychiatric Association (WPA). 2017 Oct;16(3):287-98. doi: 10.1002/wps.20472. PMID: 28941113.
7. Torok M, Han J, Baker S, et al. Suicide prevention using self-guided digital interventions: a systematic review and meta-analysis of randomised controlled trials. The Lancet Digital health. 2020 Jan;2(1):e25-e36. doi: 10.1016/s2589-7500(19)30199-2. PMID: 33328037.
8. Bates DW, Landman A, Levine DM. Health Apps and Health Policy: What Is Needed? Jama. 2018 Nov 20;320(19):1975-6. doi: 10.1001/jama.2018.14378. PMID: 30326025.
9. Parker L, Halter V, Karliychuk T, et al. How private is your mental health app data? An empirical study of mental health app privacy policies and practices. International journal of law and psychiatry. 2019 May-Jun;64:198-204. doi: 10.1016/j.ijlp.2019.04.002. PMID: 31122630.
10. Huckvale K, Torous J, Larsen ME. Assessment of the Data Sharing and Privacy Practices of Smartphone Apps for Depression and Smoking Cessation. JAMA network open. 2019 Apr 5;2(4):e192542. doi: 10.1001/jamanetworkopen.2019.2542. PMID: 31002321.
11. Clarke J, Draper S. Intermittent mindfulness practice can be beneficial, and daily practice can be harmful. An in depth, mixed methods study of the "Calm" app's (mostly positive) effects. Internet interventions. 2020 Mar;19:100293. doi: 10.1016/j.invent.2019.100293. PMID: 31890639.
12. National Institute for Health and Care Excellence. Evidence standards framework for digital health technologies. <https://www.nice.org.uk/about/what-we-do/our-programmes/evidence-standards-framework-for-digital-health-technologies>.
13. GOV.UK-Department of Health & Social Care. Guidance-A guide to good practice for digital and data-driven health technologies. <https://www.gov.uk/government/publications/code-of-conduct-for-data-driven-health-and-care-technology/initial-code-of-conduct-for-data-driven-health-and-care-technology>.

## V. Summary of Protocol Amendments

| Date                            | Section  | Original Protocol   | Revised Protocol  | Rationale  |
|---------------------------------|--|---|---|--|
| October 15 <sup>th</sup> , 2021 | Title  | Evaluation of Mental Health Applications  | Evaluation of Mental Health Mobile Applications   | The term “mobile” was added to the title for clarity.  |
| October 15 <sup>th</sup> , 2021 | E. Application of framework.<br><br>(Second - Third paragraph) | <p>Based on the a priori criteria, apps will be screened for inclusion until the target number of apps (3-5) in each bucket is reached, not to exceed a total 100 apps.</p> <p>The included apps will be downloaded. Team members will first conduct pilot testing of framework and instructions by applying the framework to 10-20 apps. This will be completed for the draft report. During peer review, the framework will be applied to the remaining apps. At least 2 members of the research team will apply the framework to 20% of the included apps. Depending on the structure of the final framework, apps will receive a score of “1” to indicate the presence of a certain feature, or a “0” to indicate the absence of the feature. Interrater reliability between the two reviewers will be assessed and any discrepancies in the application of the framework will be recorded, discussed with the rest of the research team, and definitions will be further clarified. We will then continue applying the framework to the rest of the apps, with one</p> | <p>For the draft report, team members applied the framework to 10 apps. During the peer review, the framework was applied to 17 more apps by two members of the team. An additional 18 apps will be evaluated by two team members during public review. Thus, the framework, and guidance, will be refined through dual review of a total of 45 apps.</p> | <p>The originally planned single review of apps would not add much to the refinement of framework nor provide substantive information to discuss generalizability of results for mental health apps. We changed process to dual review for all apps to further refine framework.</p> |



|  |  |  |  |  |
|--|--|--|--|--|
|  |  | <p>person applying the framework to each app. We will discuss the process of applying the framework to generate further refinements of the framework and instructions.</p> |  |  |
|--|--|--|--|--|

## VI. Key Informants

Within the Technical Brief process, Key Informants serve as a resource to offer insight into the clinical context of the technology/intervention, how it works, how it is currently used or might be used, and which features may be important from a patient or policy standpoint. They may include clinical experts, patients, manufacturers, researchers, payers, or other perspectives, depending on the technology/intervention in question. Differing viewpoints are expected, and all statements are crosschecked against available literature and statements from other Key Informants. Information gained from Key Informant interviews is identified as such in the report. Key Informants do not do analysis of any kind nor contribute to the writing of the report and will not review the report, except as given the opportunity to do so through the public review mechanism.

Key Informants must disclose any financial conflicts of interest greater than \$5,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals are invited to serve as Key Informants and those who present with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

## VII. Peer Reviewers

Peer reviewers are invited to provide written comments on the draft report based on their clinical, content, or methodologic expertise. Peer review comments on the draft report are considered by the EPC in preparation of the final report. Peer reviewers do not participate in writing or editing of the final report or other products. The synthesis of the scientific literature presented in the final report does not necessarily represent the views of individual reviewers. The dispositions of the peer review comments are documented and may be published three months after the publication of the Evidence report.

Potential Reviewers must disclose any financial conflicts of interest greater than \$5,000 and any other relevant business or professional conflicts of interest. Invited Peer Reviewers may not have any financial conflict of interest greater than \$5,000. Peer reviewers who disclose potential business or professional conflicts of interest may submit comments on draft reports through the public comment mechanism.

## **VIII. EPC Team Disclosures**

EPC core team members must disclose any financial conflicts of interest greater than \$1,000 and any other relevant business or professional conflicts of interest. Related financial conflicts of interest that cumulatively total greater than \$1,000 will usually disqualify EPC core team investigators.

## **IX. Role of the Funder**

This project was funded under Contract No. <#> from the Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services. The AHRQ Task Order Officer reviewed contract deliverables for adherence to contract requirements and quality. The authors of this report are responsible for its content. Statements in the report should not be construed as endorsement by the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.

## Appendix A. Preliminary Search Strategy

|  | <b>Search terms</b>   |
|--|---|
| <b>Application</b>                                 | mobile health [mh] OR mobile device [mh] OR tablets [mh] OR mobile applications [mh] OR “smartphone” [mh] OR mobile phone [mh] OR “mHealth”[tiab] OR “m-health”[tiab] OR “mobile health” [tiab] OR “mobile device”[tiab] OR tablet[tiab] OR “mobile app*”[tiab] OR smartphone [tiab] OR “mobile phone”[tiab] OR “digital medicine”[tiab] OR “digital health”[tiab] OR “ehealth”[tiab] OR “wireless system”[tiab] OR “electronic adherence”[tiab] OR wearable [tiab]   |
| <b>Mental health</b>                               | “Mental health”[mh] OR depression[mh] OR anxiety [mh] OR psychosis [mh] OR post-traumatic stress disorder[mh] OR “Substance-Related Disorders”[mh] OR sleep disorders[mh] OR suicide [mh] OR schizophrenia [tiab] OR Bipolar Disorder [mh] OR “mental health”[tiab] OR depression[tiab] OR anxiety [tiab] OR bipolar [tiab] OR psychosis [tiab] OR “stress disorder” [tiab] OR “substance use” [tiab] OR “sleep disorders” [tiab] OR “suicidal behaviors” [tiab] OR suicide [tiab] OR “smoking cessation”[mh] OR smoking [tiab] |
| <b>Framework</b>                                   | Framework [tiab]  |
| <b>Application AND Mental health AND Framework</b> | <b>361 citations</b>  |