



Technical Brief
Number 41

Evaluation of Mental Health Mobile Applications



Technical Brief

Number 41

Evaluation of Mental Health Mobile Applications

Prepared for:

Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
5600 Fishers Lane
Rockville, MD 20857
www.ahrq.gov

Contract No. 75Q80120D00003

Prepared by:

Johns Hopkins University Evidence-based Practice Center
Baltimore, MD

Investigators:

Smisha Agarwal, Ph.D.
Madhu Jalan, M.S., M.B.A.
Holly C. Wilcox, Ph.D.
Ritu Sharma, B.Sc.
Rachel Hill, B.A.
Emily Pantalone, B.A.
Johannes Thrul, Ph.D.
Jacob C. Rainey, M.P.H.
Karen A. Robinson, Ph.D.

AHRQ Publication No. 22-EHC016

May 2022

This report is based on research conducted by the Johns Hopkins University Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. 75Q80120D00003). The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

None of the investigators have any affiliations or financial involvement that conflicts with the material presented in this report.

The information in this report is intended to help healthcare decisionmakers—patients and clinicians, health system leaders, and policymakers, among others—make well-informed decisions and thereby improve the quality of healthcare services. This report is not intended to be a substitute for the application of clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information, i.e., in the context of available resources and circumstances presented by individual patients.

This report is made available to the public under the terms of a licensing agreement between the author and the Agency for Healthcare Research and Quality. This report may be used and reprinted without permission except those copyrighted materials that are clearly noted in the report. Further reproduction of those copyrighted materials is prohibited without the express permission of copyright holders.

AHRQ or U.S. Department of Health and Human Services endorsement of any derivative products that may be developed from this report, such as clinical practice guidelines, other quality enhancement tools, or reimbursement or coverage policies, may not be stated or implied.

AHRQ appreciates appropriate acknowledgment and citation of its work. Suggested language for acknowledgment: This work was based on an evidence report, Evaluation of Mental Health Mobile Applications, by the Evidence-based Practice Center Program at the Agency for Healthcare Research and Quality (AHRQ).

Suggested citation: Agarwal S, Jalan M, Wilcox HC, Sharma R, Hill R, Pantalone E, Thrul J, Rainey JC, Robinson KA. Evaluation of Mental Health Mobile Applications. Technical Brief 41. (Prepared by the Johns Hopkins University Evidence-based Practice Center under Contract No. 75Q80120D00003.) AHRQ Publication No. 22-EHC016. Rockville, MD: Agency for Healthcare Research and Quality; May 2022. DOI: <https://doi.org/10.23970/AHRQEPCTB41>. Posted final reports are located on the Effective Health Care Program [search page](#).

Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of evidence reports and technology assessments to assist public- and private-sector organizations in their efforts to improve the quality of healthcare in the United States. The reports and assessments provide organizations with comprehensive, science-based information on common, costly medical conditions and new healthcare technologies and strategies. The EPCs systematically review the relevant scientific literature on topics assigned to them by AHRQ and conduct additional analyses when appropriate prior to developing their reports and assessments.

This EPC evidence report is a Technical Brief. A Technical Brief is a rapid report, typically on an emerging medical technology, strategy, or intervention. It provides an overview of key issues related to the intervention—for example, current indications, relevant patient populations and subgroups of interest, outcomes measured, and contextual factors that may affect decisions regarding the intervention. Although Technical Briefs generally focus on interventions for which there are limited published data and too few completed protocol-driven studies to support definitive conclusions, the decision to request a Technical Brief is not solely based on the availability of clinical studies. The goals of the Technical Brief are to provide an early objective description of the state of the science, a potential framework for assessing the applications and implications of the intervention, a summary of ongoing research, and information on future research needs. In particular, through the Technical Brief, AHRQ hopes to gain insight on the appropriate conceptual framework and critical issues that will inform future research.

AHRQ expects that the EPC evidence reports and technology assessments will inform individual health plans, providers, and purchasers as well as the healthcare system as a whole by providing important information to help improve healthcare quality.

If you have comments on this Technical Brief, they may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 5600 Fishers Lane, Rockville, MD 20857, or by email to epc@ahrq.hhs.gov.

Robert Otto Valdez, Ph.D., M.H.S.A.
Director
Agency for Healthcare Research and Quality

Arlene S. Bierman, M.D., M.S.
Director
Center for Evidence and Practice
Improvement
Agency for Healthcare Research and Quality

Craig A. Umscheid, M.D., M.S.
Director
Evidence-based Practice Center Program
Center for Evidence and Practice
Improvement
Agency for Healthcare Research and Quality

David W. Niebuhr, M.D., M.P.H., M.Sc.
Task Order Officer
Center for Evidence and Practice
Improvement
Agency for Healthcare Research and Quality

Acknowledgments

The authors gratefully acknowledge the following individuals for their guidance: Hadi Kharrazi, M.D., Ph.D., Harold P. Lehmann, M.D. Ph.D., Alain B. Labrique, Ph.D., M.H.S., M.S., and Joseph Ali, J.D. We thank Jina Ryu, B.A., Mehek Bapna, B.A., and Pritika Parmer, B.A., for their help with abstraction. We also thank Jeanette Edelstein, M.A., for her services in copyediting this report.

Key Informants

In designing the study questions, the EPC consulted a panel of Key Informants who represent subject experts and end-users of research. Key Informant input can inform key issues related to the topic of the Technical Brief. Key Informants are not involved in the analysis of the evidence or the writing of the report. Therefore, in the end, study questions, design, methodological approaches and/or conclusions do not necessarily represent the views of individual Key Informants.

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 role as end-users, individuals with potential conflicts may be retained. The Task Order Officer (TOO) and the EPC work to balance, manage, or mitigate any conflicts of interest.

The list of Key Informants who provided input to this report follows:

Anita Bajaj, M.D., M.P.H.*
Food and Drug Administration
Silver Spring, MD

Kate Farinholt
NAMI Maryland
Columbia, MD

Lori M. Hilt, Ph.D.*
Lawrence University
Appleton, WI

Beth Jaworski, Ph.D.
U.S. Department of Veterans Affairs
Menlo Park, CA

Adam Kaplin, M.D., Ph.D.*
MyMD Pharmaceuticals
Baltimore, MD

Oliver J. Kim, J.D., L.L.M.*
Cambia Health Solutions
Portland, OR

Lauren Lowenstein, M.P.H.*
Centers for Medicare & Medicaid Services
Baltimore, MD

Douglas Olson, M.D.
Centers for Medicare & Medicaid Services
Baltimore, MD

Brian Suffoletto, M.D., M.S.*
Stanford University
Stanford, CA

Michelle Tye, Ph.D.*
Black Dog Institute, University of New South Wales
Randwick, Australia

Stephen Warnick Jr., M.D.*
Oakland University William Beaumont School of Medicine
Rochester, MI

*Provided input on Draft Report.

Peer Reviewers

Prior to publication of the final evidence report, EPCs sought input from independent Peer Reviewers without financial conflicts of interest. However, the conclusions and synthesis of the scientific literature presented in this report do not necessarily represent the views of individual reviewers.

Peer Reviewers 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 with potential nonfinancial conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential nonfinancial conflicts of interest identified.

The list of Peer Reviewers follows:

Trina Histon, Ph.D.
Kaiser Permanente
Oakland, CA

Charmain F. Jackman, Ph.D.
InnoPsych
Boston, MA

Dara H. Sorkin, Ph.D.
University of California
Irvine, CA

Evaluation of Mental Health Mobile Applications

Structured Abstract

Background. Mental health mobile applications (apps) have the potential to expand the provision of mental health and wellness services to traditionally underserved populations. There is a lack of guidance on how to choose wisely from the thousands of mental health apps without clear evidence of safety, efficacy, and consumer protections.

Purpose. This Technical Brief proposes a framework to assess mental health mobile applications with the aim to facilitate selection of apps. The results of applying the framework will yield summary statements on the strengths and limitations of the apps and are intended for use by providers and patients/caregivers.

Methods. We reviewed systematic reviews of mental health apps and reviewed published and gray literature on mental health app frameworks, and we conducted four Key Informant group discussions to identify gaps in existing mental health frameworks and key framework criteria. These reviews and discussions informed the development of a draft framework to assess mental health apps. Iterative testing and refinement of the framework was done in seven successive rounds through double application of the framework to a total of 45 apps. Items in the framework with an interrater reliability under 90 percent were discussed among the evaluation team for revisions of the framework or guidance.

Findings. Our review of the existing frameworks identified gaps in the assessment of risks that users may face from apps, such as privacy and security disclosures and regulatory safeguards to protect the users. Key Informant discussions identified priority criteria to include in the framework, including safety and efficacy of mental health apps. We developed the Framework to Assist Stakeholders in Technology Evaluation for Recovery (FASTER) to Mental Health and Wellness and it comprises three sections: Section 1. Risks and Mitigation Strategies, assesses the integrity and risk profile of the app; Section 2. Function, focuses on descriptive aspects related to accessibility, costs, organizational credibility, evidence and clinical foundation, privacy/security, usability, functions for remote monitoring of the user, access to crisis services, and artificial intelligence (AI); and Section 3. Mental Health App Features, focuses on specific mental health app features, such as journaling and mood tracking.

Conclusion. FASTER may be used to help appraise and select mental health mobile apps. Future application, testing, and refinements may be required to determine the framework's suitability and reliability across multiple mental health conditions, as well as to account for the rapidly expanding applications of AI, gamification, and other new technology approaches.

Contents

Executive Summary	1
Introduction	1
Background.....	1
Objectives of This Technical Brief.....	3
Guiding Questions	4
Methods	5
Identification and Abstraction of Existing Frameworks.....	5
Identification of Frameworks To Evaluate Health Apps.....	5
Abstraction of Items From Existing Frameworks	6
Identification of Gaps	6
Key Informant Interviews: Round One	6
Analyses of the Existing Frameworks.....	6
Development of the New Framework.....	6
Consolidation of Items From Existing Frameworks	6
Development of Criteria for New Thematic Areas	7
Refinement of the New Framework.....	9
Key Informant Interviews: Round Two	9
Testing and Refinement.....	10
Peer Review and Public Commentary	12
Findings	13
Gaps in Existing Frameworks.....	13
Risk and Credibility.....	13
Vulnerable Populations	13
Accessibility Features.....	13
Cultural Sensitivity	13
Artificial Intelligence.....	14
Mental Health Features and Function.....	14
Results From Key Informant Interviews	14
Key Informant Interviews Round One	14
Key Informant Interviews Round Two.....	15
Testing and Refinement of the Framework During Pre-Pilot and Pilot Rounds	17
Results of the App Search Process on 42matters	17
Testing and Refinement.....	19
Summary of Changes Made to the Framework During Pilot Testing Rounds	20
Development of a Pre-Assessment Questionnaire	20
Endorsements and Usage.....	20
Assessment of App Risk Based on “Recommended Age”.....	21
“Special” to “Vulnerable” Populations	21
Privacy Policy and Terms of Agreement.....	21
Cultural Competence	22
Changes to Usability Criteria	22
Interaction With a Coach or Healthcare Professional	22
Categorization of Mental Health Features.....	22
Acceptable Research	22

Subjectivity Versus Objectivity in Question Responses	23
Overview of Results of Application of FASTER.....	23
Overview of the FASTER to Mental Health and Wellness Framework.....	23
Target Audience and Intended Use	24
Administrative Questionnaire.....	25
Section 1: Risks and Mitigation Strategies.....	26
Section 2: Function.....	30
Section 3: Mental Health App Features.....	37
Post-Administrative Questionnaire	40
Summary and Implications.....	42
Next Steps	42
References.....	45

Tables

Table A. Summary of sections and questions In FASTER.....	2
Table 1. Categorization of mental health disorders	10
Table 2. Summary of the interrater reliability in each round and number of changes made to framework to improve interrater reliability	20
Table 3. Administrative questionnaire.....	25
Table 4. Section 1: Risk and Mitigation Strategies.....	26
Table 5a. Step 1: Preliminary Safety Assessment for Risk Levels 2 and 3	27
Table 5b. Step 2: Safety Assessment for Vulnerable Population	28
Table 5c. Step 3: Safety Assessment	28
Table 6. App integrity levels.....	28
Table 7. Risk levels.....	29
Table 8. Section 2: Function	30
Table 9. Section 3: Mental Health App Features.....	37
Table 10. Post-Administrative Questionnaire.....	40

Boxes

Box 1. Key Informants' expertise, with number of participants.....	14
Box 2. Ideal components/characteristics of a framework: summary from Key Informant interviews.....	16

Figures

Figure 1. Framework development process.....	5
Figure 2. Mental health mobile apps screening process	18
Figure 3. Framework to Assist Stakeholders in Technology Evaluation for Recovery (FASTER) to Mental Health and H.....	24
Figure 4. Categories of questions in each section.....	24

Appendixes

Appendix A. List of Existing Frameworks
Appendix B. Framework Criteria
Appendix C. Modifications in uMARS Questions
Appendix D. Pilot Rounds Results

Appendix E. Framework Training Guide: Framework to Assist Stakeholders in Technology Evaluation for Recovery (FASTER) to Mental Health and Wellness Framework Assessment
Appendix F. Framework
Appendix G. Framework Assessment Examples

Executive Summary

Main Points

- We developed and pilot tested a framework to assess mental health mobile applications (apps).
- The Framework to Assist Stakeholders in Technology Evaluation for Recovery (FASTER) to Mental Health and Wellness assesses the risk/safety, technical functionality, and mental health features of apps.
- FASTER can be used by advocacy organizations, payers, healthcare systems, and others to inform selection of mental health mobile apps.

Background and Purpose

People suffering from mental or behavioral health conditions may not seek or receive care owing to stigma, provider shortages, long wait times, lack of proximity to providers, or other accessibility barriers. Mental health mobile applications (apps) may help to address this gap. However, the potential app privacy/security and safety concerns are even more salient in a vulnerable population of users, and thus the decisional dilemma is, “How can consumers, family members and peer supports, providers, and health systems select safe and effective mental health and wellness apps best suited to their needs?” The aim of this Technical Brief was to develop a framework to assess mental health apps for users across age groups, for different mental health symptoms and disorders, and for general mental wellness.

Methods

We reviewed existing frameworks, published review articles on frameworks, and the gray literature. We also conducted Key Informant interviews to identify elements to inform the development of a framework to assess the safety and effectiveness of mental health apps. A second group of Key Informants provided feedback on the initial framework. Iterative testing and refinement of the framework was done in seven successive rounds (1 pre-pilot round and 6 pilot rounds) through double application of the framework to a total of 45 apps.

Results

We found that the existing frameworks have a heavy emphasis on technology; have limited information about app features specific to mental health; and do not account for potential risks, including those posed by advances in machine learning and artificial intelligence (AI).

The FASTER to Mental Health and Wellness framework expands on existing frameworks to address the identified gaps, especially in the area of risk assessment and use of features specific to mental health. It comprises three sections with an initial and concluding set of administrative questions: Section 1: Risks and Mitigation Strategies, assesses the integrity and risk profile of the app; Section 2: Function, focuses on descriptive aspects related to accessibility, costs, organizational credibility, evidence and clinical foundation, privacy and security, usability, functions for remote monitoring of the user, informed consent, cultural competence, access to crisis services, and AI; and Section 3: Mental Health App Features, focuses on specific features and functions of the app that could align with and facilitate therapeutic goals. Within each of these sections, there are a series of items related to the assessment of specific categories that

were considered critical based upon the literature review and Key Informant interviews. A summary of FASTER, including the number of questions and response type information in each section, is provided in Table A.

Table A. Summary of sections and questions In FASTER

Section	Category	Number of Questions	Response Type
Administrative Questionnaire	Administrative questionnaire	14	Free text
1. Risk and Mitigation Strategies	App Integrity	4	Multiple choice
	Risk Assessment	3	Multiple choice
	Evidence	2	Multiple choice
	Linkage to Care	1	Multiple choice
	Access to Crisis Resources	2	Multiple choice
2. Function	Accessibility Features	2	Multiple choice
	App Information	6	Free text
	Costs	4	Multiple choice
	Organizational Credibility	2	Multiple choice
	Evidence and Clinical Foundation	1	Multiple choice
	Privacy and Security	5	Multiple choice
	Informed Consent	2	Multiple choice
	Cultural Competence	4	Multiple choice
	Usability	15	Multiple choice
	Remote Monitoring	3	Multiple choice
	Access to Crisis Response Services	1	Multiple choice
	Artificial Intelligence (AI)	3	Multiple choice, free text
	3. Mental Health App Features	Mental health app features	3 (1 question has 7 sub questions)
Mental health functionality		2 (1 question has 15 sub-questions)	Multiple choice, free text
Post Administrative Questionnaire	Post administrative questionnaire	10	Multiple choice, free text

FASTER: Framework to Assist Stakeholders in Technology Evaluation for Recovery

Strengths and Limitations

The FASTER to Mental Health and Wellness framework is aimed at facilitating the use of apps for mental health support and recovery through standardized evaluation, screening, and classification of apps. Several of the criteria have been extracted from existing frameworks in the app evaluation and mental health domain. However, we identified several gaps in the existing frameworks and addressed them through further prioritization of criteria, the addition of criteria to assess risks and safety of the apps, assessment of specific mental health app features that could facilitate therapeutic goals, and assessment of the use of AI and other customized engagement approaches.

We acknowledge that apps targeted at specific disorders may benefit from frameworks specific to those disorders; however, we also recognize that developing frameworks that are specific to disease areas is time and resource intensive. As a next step to facilitating uptake and adoption, it would be valuable to gather user (e.g., provider, patient) input on the applicability of specific components of the FASTER framework in guiding decisions about the use of mental health apps.

Implications and Conclusions

We envisage a range of possible applications of this framework. First, the framework can be used by mental health organizations and mental health advocacy agencies to provide a curated list or library of safe and effective mental health apps. Such a library could be used by consumers, family members, peer supporters, and health care providers to review and select apps as a resource for patients. Second, the framework can be leveraged by employee health plans, health system leaders, public and private insurance providers, and other entities to review and provide guidance for apps relevant to their members. Lastly, the framework can be used by app developers as guidance to promote transparency in communication about the potential benefits, risks, and evidence to support their apps.

Introduction

Background

Among adults 18 years of age and older in the United States, the prevalence of ‘mental disorders in the past year’ increased from 17.7 percent (or 39.8 million people) in 2008 to 20.6 percent (or 51.5 million people) in 2019.¹ Of this latter population, 26.0 percent (or 13.3 million people) perceived an unmet need for mental health services. The most common reason given for unmet need was affordability, with other barriers including mental health stigma, provider shortages, and wait time.²

According to the World Health Organization, “mental health is a state of well-being in which an individual realizes his or her own abilities, can cope with the normal stresses of life, can work productively and is able to make a contribution to his or her community.”³ Mobile health apps (mHealth) have great potential to provide much needed access and equity in mental health care and wellness. Mobile health apps are applications that run on smartphones or tablet devices and contain content to facilitate health and wellness. Mental health apps, a subset of mobile health apps, are being used to support assessment, diagnosis, treatment, and management of mental health, as well as to provide wellness support through meditation and mindfulness. Some mental health apps can provide diagnostic support or assist in the diagnostic pathway by improving time to diagnosis, for example by offering automated standardized mental health assessments. Apps might also facilitate treatment for certain mental health conditions and provide therapeutic support. Firth et al. found in their meta-analysis that apps targeting depression through self-management significantly reduced symptoms of depression when compared to controls.^{4,5} The use of mental health apps can also reduce anxiety compared with controls.⁶

Mental health apps provide patients with relevant health information and help enhance quality of care by supporting patients through access to knowledge and other skill-building resources. At an individual level, mental health apps can be a resource that provides psychoeducation and medication management, skill building, and symptom tracking thereby amplifying the benefits of existing therapy. They also add the convenience of accessing mental health care remotely and may provide more privacy than in-person mental health visits, potentially avoiding stigma associated with seeking mental health care. Additionally, they may further improve access by lowering the cost of care.⁷ Mental health apps can be readily disseminated with minimal need for staff training or resource investment, and thereby reduce the burden on healthcare settings. Furthermore, technology-enhanced healthcare approaches are increasingly being reimbursed by insurers,⁸ making such apps more likely to be integrated into the toolkits of healthcare providers. Improved care coordination, more touchpoints to enhance treatment with a provider, and addressing patient-centered care is another potential benefit of mental health apps.⁷ Some apps deliver greater efficiency by providing feedback loops and assessments for providers in between appointments and also deliver round-the-clock support and direct access to crisis lines for patients.⁹

Significant racial and ethnic disparities exist in access to and use of mental health care services.^{10,11} While some studies explain that disparities exist owing to a lack of access to care (including a lack of insurance coverage), others, when comparing similar sociodemographic factors, suggest that the differences may be more related to mistrust of traditional mental health services, discrimination, language, and cultural factors.¹² Black and Asian individuals are screened for depression at a lower rate compared with white individuals and, for those that receive screening, Black individuals, Latino males, and Asian individuals receive mental health

care at a lower rate compared with their white counterparts.¹³ Disparities between urban and rural populations' access to mental health services exist as well. Up to 65 percent of non-urban counties lack psychiatrists, while more than 60 percent of Americans in rural areas live in a Mental Health Provider Shortage Area.¹⁴ Apps could help address disparities in mental health services and provide a gateway to care for those who do not trust health professionals or have not benefitted from traditional mental health services. Many mental health apps are affordable, offer a way to reach rural populations and traditionally hard-to-reach groups, and can reduce patients' feelings of stigmatization or discomfort with traditional treatment.¹⁵

The market for mental and behavioral health is booming, turbocharged by funding through private and venture-backed investors. In 2020, \$2.4 billion was raised for startups working in mental health, equivalent to 17 percent of all funding for digital health in 2020.¹⁶ Segmenting the space into only apps, \$1.2 billion were raised in 2021, which is a 50 percent increase compared with the prior year.¹⁷ Some driving factors for funding may be the huge increase in demand for digital mental health services owing to isolation, job loss, economic and financial hardships and other detrimental effects of the COVID-19 pandemic. For the first time, during the pandemic, the Food and Drug Administration (FDA) began approving digital mental health therapeutics solutions and federal regulatory requirements allowed more flexibility for telehealth.¹⁸

Mobile apps which fall into the Software as a Medical Device (SaMD) category for the FDA follow the same regulatory guidelines as other medical devices with three classifications for approval. Class I medical devices have a low to moderate risk to the patient and/or user and a premarket notification application and FDA clearance is not required before marketing the device in the United States. However, the manufacturer is required to register and list their product with the FDA. Class II medical devices have a moderate to high risk to the patient and/or user. For example, a computerized behavioral therapy device for psychiatric disorders regulated under 21 CFR 882.5801 is a prescription-only device intended to provide a computerized version of condition-specific behavioral therapy as an adjunct to clinician supervised outpatient treatment to patients with psychiatric conditions. The FDA may remove from the market any Class II and Class III devices that are determined to be fraudulent or potentially harmful. An example of a such a circumstance includes apps with treatment claims for specific psychiatric conditions where the underlying condition may require an urgent or immediate clinical intervention and for which the delay of the intervention may pose significant harm to the patient (e.g., treatment of suicidality). Class III medical devices are those devices that have a high risk to the patient and/or user. These devices usually sustain or support life, or present potential unreasonable risk of illness or injury.

Not all mental health apps are considered SaMD. While there are about 20,000 mental health apps on the market in the Apple App Store or Google Play Store, only five of them have FDA approval.¹⁹ The first prescription digital therapeutic for mental health, which launched in 2017, was Pear Therapeutics' reSET therapy, which is powered through an app.²⁰ The FDA has taken a "hands-off" approach to the regulation of low risk, general wellness apps. The FDA has also introduced a Software Precertification (Pre-Cert) Pilot Program that provides streamlined regulatory oversight of software-based medical devices developed by certain manufacturers that have consistently demonstrated quality and organizational excellence. By pre-certifying organizations instead of specific products, the program aims to reduce burdensome regulatory oversight for these organizations as they develop high quality SaMD products.^{21, 22}

Rapid proliferation of health apps has resulted in both haphazard and sub-optimal use of these apps, with potential dangers to patients and end-users because the health system and

providers may not be well-versed enough regarding the purpose, safety, and efficacy of these apps to recommend or prescribe them. There is no existing gold standard tool for informing the selection of mental health apps for patients and clinicians. “Prescribing” an app in a medical setting is challenging owing to the limited evidence on the efficacy of the apps, as well as emerging concerns about the usability, privacy, and safety risks.²³

Apps that target individuals who may be vulnerable to scams or fraud owing to their mental health impairment can pose a serious risk.²³ Mental health apps may pose a high risk to users with serious or acute mental disorders, such as untreated psychosis and suicide ideation, especially if the apps do not provide easy access to crisis lines in case of emergency. According to Martinengo et al., there have been more than 2 million downloads of mental health apps with inaccurate or without suicide crisis phone numbers.²⁴ Privacy and security, particularly related to private health information, is also a concern for mental health apps. Dehling et al. ranked the level of damage to users owing to privacy infringement and found that 95 percent of apps pose at least some threat.²⁵ Additionally, mental health apps have possible data security risks and the risk of misinformation.²⁶ The regulatory framework has not kept up with the proliferation of mental health apps. One major risk involves data security such as data breaches resulting in exposures to various risks such as scams, identity theft, and exposure to marketing. Mental health app content could provide harmful advice, disseminate false information, or could have iatrogenic impacts due to insensitive homophobic or racist content. Mental health providers should consider liability when recommending an app to a patient.

There are many frameworks that are being used to evaluate digital health apps, including those that focus on mental health apps.²⁷⁻³¹ Several of these frameworks are being used by advocacy agencies and online health resource platforms to recommend apps to interested users. However, most existing frameworks are geared towards evaluating specific aspects of health apps (e.g., usability), and do not adequately reflect concerns around assessment of risks posed by the apps, as well as recent advancements in artificial intelligence, and their use by apps for automating certain diagnostic (e.g., Ada) or counseling protocols (e.g., Woebot, Replika).

Objectives of This Technical Brief

Given the uncertain evidence-base for most mental health apps coupled with potential app safety and privacy concerns in a more vulnerable population of users, the decisional dilemma is “How should consumers, family members and peer supports, and providers and health systems select mental health and wellness mobile apps?” The aim of this Technical Brief is to develop a framework to assess the safety and effectiveness of mental health and wellness apps for users across age groups. This Technical Brief addresses only mobile health apps that can be downloaded on a smartphone and does not include pure text-based apps, wearable devices, or general telehealth and telemedicine apps.

Guiding Questions

This Technical Brief was guided by the following questions:

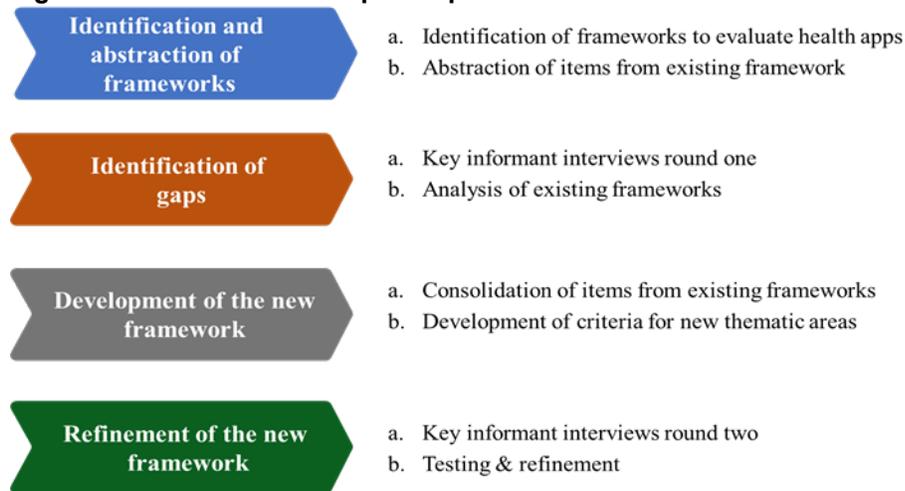
1. What characteristics and minimal standards of available mental health mobile apps need to be analyzed to assess the appropriateness (to various stakeholders) and effectiveness of available apps, to include but not be limited to:
 - Accessibility, including ease of use, health literacy, 508 compliance,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

Methods

The methods for this Technical Brief follow the Content and Procedures Guide for the Evidence-based Practice Center Program. The protocol was posted on the Agency for Healthcare Research and Quality (AHRQ) Effective Health Care website (<https://effectivehealthcare.ahrq.gov/products/mental-health-apps/protocol>). The Technical Brief is organized by two Guiding Questions that address the development of a framework for assessing mental health apps.

Figure 1 displays the process of developing the framework.

Figure 1. Framework development process



Identification and Abstraction of Existing Frameworks

Identification of Frameworks To Evaluate Health Apps

We identified a number of frameworks that focused on the assessment of general health apps, as well as mental health apps. Our review of existing frameworks was guided by our knowledge of frameworks, a literature search for mental health app assessment frameworks using PubMed[®], and the identification of additional sources through a reference review of included documents. Examples of frameworks for the assessment of general health apps include MyHealthApps[™],³² Healthy Living Apps Guide,³³ Digital Technology Assessment Framework from National Health Service UK,³⁴ and Digital Therapeutics Alliance.³⁵ Examples of frameworks that are focused on mental health app assessment include One Mind Psyber Guide,²⁷ American Psychiatric Association Initiative,²⁸ Kaiser Permanente,²⁹ VeryWellMind,³⁰ and HealthNavigator.³¹ Other notable frameworks reviewed include M-Health Index and Navigation Database (MIND) and the end-user version of the Mobile Application Rating Scale (uMARS). MIND is an operational and flexible framework based on the American Psychiatric Association App Assessment Framework, which includes 105 questions that have been harmonized from 79 frameworks.³⁶ The end-user version of uMARS provides a comprehensive set of questions about app engagement and usability.³⁷

We also conducted a search for documents related to the regulation of health apps, including by the Food and Drug Administration (FDA), as well as by Germany's Federal Institute for Drugs and Medical Devices and the United Kingdom's National Institute for Health and Care

Excellence (NICE) Framework.^{38, 39} Finally, we identified relevant systematic reviews on the efficacy and limitations of typical features found in mental health apps, such as mindfulness, mood, and symptom trackers, journaling, social and peer interaction, psychoeducation, and skill-building, to better understand how they might be relevant to a framework to assess mental health apps.^{24, 40-44}

Abstraction of Items From Existing Frameworks

We abstracted items from 11 existing frameworks. (See Appendix A for a list of frameworks.) This process yielded 300 items/questions. There were significant similarities between frameworks and various items typically aimed to capture similar concepts, such as aspects of credibility, safety, available features to align with therapeutic goals, accessibility features for those with disabilities, clinical evidence foundation, interoperability, usability, efficacy, and data privacy and security and were therefore categorized under common themes. Items that were clear and easier to operationalize were retained. Those that were too technical to be operationalized by individuals without specialized training, were confusing and unclear, or required extensive research to answer were simplified or removed.

Identification of Gaps

To identify areas for further development of an effective framework to evaluate mental health apps, we conducted Key Informant (KI) interviews to identify priorities from various stakeholder perspectives and analyzed categories that were covered in the existing frameworks.

Key Informant Interviews: Round One

We selected KIs to represent a range of expertise and perspective from stakeholders. In the first round, we conducted interviews with stakeholders representing family members of those living with mental disorders; clinicians with a background in mental health, primary health care, and emergency medicine; and payors. We asked these stakeholders to provide guidance on their perceptions and experience with mental health apps and the essential features and omissions in existing frameworks with which they had familiarity. The family/patient representative was identified from one of the largest mental health advocacy organizations, National Alliance on Mental Illness (NAMI).

Analyses of the Existing Frameworks

We reviewed the categories covered by the existing app evaluation frameworks, as well as other literature highlighting priority considerations for such frameworks. This literature included documents on regulatory and safety considerations for medical software. Based on this assessment, and feedback from the interviews with the KIs, we identified key gaps in existing frameworks, which are described in further detail in the Results section below.⁴⁵⁻⁴⁷

Development of the New Framework

Consolidation of Items From Existing Frameworks

We systematically selected, adapted, and where necessary, modified the abstracted criteria from the frameworks. (See Appendixes B and C.) To the extent possible, we leveraged existing

criteria that have been proposed and used by other frameworks. Where necessary, we simplified the abstracted criteria and standardized the language for clarity. We were guided by the following principles in the abstraction and revision of the criteria: (1) the question posed by the criteria should be relatively easy to answer based on a review of the app developer website and through the process of downloading the app (it should not require a systematic literature search or engagement with the app developer); (2) anyone with some knowledge of technology and mental health should be able to use the framework, but evaluators do not need to be experts in either domain; (3) the application of the framework should be objective, with as much specificity in the criteria as possible, to increase reliability and reduce subjectivity in the evaluation process. To further facilitate the third principle, we developed a training guide for the application of the framework.

Several criteria from the MIND framework, which operationalizes the American Psychiatric Association framework, were adapted.³⁶ Several items in the uMARS framework, which has been developed and validated for the assessment of engagement, information, quality, and aesthetics, were also incorporated in our framework.³⁷ Questions on entertainment, interest, interactivity, the quantity of information, and visual information were more subjective and, hence, omitted. Questions on the credibility of the source were omitted as this issue is covered in the Risk Assessment section of our framework. Questions on the subjective quality of apps and the perceived impact of apps are covered in the post-administrative section of our framework, where the reviewer can provide their subjective assessment in free text.

Development of Criteria for New Thematic Areas

For the identified gaps that were also prioritized in the KI interviews, we developed new criteria based on peer-reviewed literature and engagement with internal experts. For the assessment of risks posed by the apps, we used normative guidance provided by government agencies such as the NICE Framework⁴⁸ and the FDA Clinical Decision Support Draft Framework and Software as a Medical Device guidance.^{22, 49, 50} For the development of criteria on the use of artificial intelligence (AI), we reviewed literature on issues of safety with the use of AI for health apps and consulted with the ethicist on our team.^{51, 52} We have described our approach to new areas of development in greater detail below.

Risk Assessment

To develop criteria on the assessment of risks of apps, we reviewed key regulatory documents, including policies from FDA's Digital Health Center of Excellence (DHCoE). DHCoE aligns and coordinates digital health work across the FDA and is charged with developing a comprehensive approach to regulation of digital health technology. The FDA has taken a "hands-off" approach towards regulating mental health apps that do not fall into the realm of "device software functions."

Device software functions may include "Software as a Medical Device (SaMD)" and "Software in a Medical Device (SiMD)." If a software function meets the definition of a device that is deployed on a mobile platform, it is referred to as a "mobile medical app." Some examples of mobile apps that are regulated by the FDA include those that use a mobile platform's built-in features (e.g., light, vibrations, and the camera) to perform medical device functions; software functions that control the operation of an implantable or body-worn medical device; and software functions that are used in active patient monitoring to analyze patient-specific medical device data.

We also reviewed FDA guidance on clinical decision support software (CDS),⁵³ including a draft document called Software as a Medical Device: Possible Framework for Risk Categorization and Corresponding Considerations which describes the risk-based approach to regulating CDS.⁵⁴ The risk categorization is divided into four categories (I, II, III, IV) based on the impact on individual health driven by the stated intent of the CDS to treat or diagnose individuals, or drive or inform clinical management. The risk categorization approach in Section 1 of our framework uses this lens by, first, assessing the level of risk the app might pose based on its stated objectives, and, second, by determining an appropriate level of evidence given the level of risk.

Informed Consent

To develop criteria related to informed consent, we studied published literature on the ethics of electronic consent and consulted with our ethics advisor.⁵⁵⁻⁵⁷ A book about digital contract tracking technologies emphasized the importance of “privacy by design”; that is, building privacy and security protections into the design of technology.⁵⁸ Consent procedures for apps usually require users to review detailed legal consent forms which may be incomprehensible. The digital contract tracking technologies stress the importance of incorporating meaningful mechanisms to obtain consent which are easily understood by the user.⁵⁸ We also looked at a simple open-source smartphone consent module that was developed by Sage Bionetworks for use in research. Its recommendation includes simple and straightforward information, deliberately organized content, multimodal learning (e.g., visual, audio, written), accessibility for disabled users, multilingual text, and engagement through interaction (e.g., swiping to navigate forward and backward).⁵⁵

App Integrity and Organizational Credibility

The trustworthiness of an app is based not just on its content, usability, and technology, but also on organizational attributes of reputation and brand.⁵⁹ Discussions with KIs highlighted the importance of a viable business model so apps could be maintained and updated regularly. Apps had the potential to cause harm that would outweigh the benefits if individuals using it suddenly found that they no longer had access to the app owing to its discontinuation or lack of updates in compliance with the latest system guidelines. Based on this guidance, we developed app integrity and organizational credibility guidelines.

Cultural Competence

The definition of cultural competence in the framework was adapted from the Health and Human Services’ definition of cultural competence:⁶⁰ “cultural competence is the integration and transformation of knowledge about individuals and groups of people into specific standards, policies, practices, and attitudes used in appropriate cultural settings to increase the quality of services; thereby producing better outcomes.”⁶⁰ For this assessment, groups with lived experiences such as pregnant teens and survivors of gender-based violence would be considered to be a specific cultural group. Research from Substance Abuse and Mental Health Services Administration (SAMHSA) shows that cultural responsiveness can improve client engagement in services, therapeutic relationships between clients and providers, and treatment retention and outcomes, and that it is critical to the reduction of disparities in behavioral health.⁶¹ Based on internal discussions, the definition of cultural competence was expanded to include groups with certain lived experiences, such as pregnant teens and survivors of gender-based violence. This

definition was further expanded to include gender sensitivity to highlight its importance in healthcare.⁴⁰

Vulnerable Populations

Discussions with the KIs highlighted the unique needs of caregivers who were responsible for either a minor or an adult with substantial impairment. In order to ensure security and safety of minors or adults with substantial impairment, there needs to be some level of monitoring by caregivers, so we added consent requirements for caregivers of vulnerable populations.

We considered defining a vulnerable adult based on the severity of mental health disorders or conditions. However, the level of disability and subsequent impact on the quality of life may vary by individuals and over time, irrespective of a mental health diagnosis, making it hard to objectively define a vulnerable adult for the purpose of this framework. Consequently, we based the definition of a vulnerable adult on their level of impairment rather than the severity of any mental health disorder diagnosis.

Mental Health Categories

Initially, we defined mental health symptoms and diagnostic categories according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-5). The DSM-5 is an authoritative source that defines and classifies mental disorders in order to improve diagnosis, treatment, and research. Guided by the symptoms and diagnostic criteria outlined in the DSM-5 and the International Classification of Diseases (ICD), we searched for systematic reviews on the efficacy on mental health apps to support a range of mental health symptoms and conditions. We reviewed recommended features in mental health apps through systematic reviews of different categories of mental health apps such as those for neurocognitive disorders, personality disorders, and sleep-wake disorders.^{24, 41-44, 62-65} Our examination of the literature showed the common features in mental health apps to be mindfulness, mood trackers, journaling, social and peer interaction, psychoeducation, and skill building. Although we initially attempted to develop a list of features of mental health apps based on the specific condition the apps address, given that features such as mindfulness and mood tracking can be useful across a range of conditions, and the evidence to guide the use of specific features for specific conditions is limited, we developed a comprehensive listing of features that mental health apps might include that can align with therapeutic goals across multiple mental health conditions, without associating these features with a specific condition.

Refinement of the New Framework

We shared the initial draft of the framework with the second round of KIs and iteratively tested and revised it through seven rounds (1 pre-pilot and 6 pilot rounds) of application of the framework to a variety of mental health apps.

Key Informant Interviews: Round Two

For the second round of KI interviews, we recruited app developers and mental health providers with app development expertise, some of whom are also mental health clinicians. To include the perspectives of KIs with expertise in developing frameworks to assess mental health apps, we worked to identify KIs with some familiarity with FDA regulations around risk stratification for apps and familiarity with existing challenges around "software as a medical

device," and privacy and security risks. The stakeholders included in the second round received a draft of the framework to review and comment on for specific recommendations.

Testing and Refinement

The draft framework was applied to a range of mental health apps and iteratively refined. Interrater reliability was calculated for each question. We used the basic interrater reliability measure, which is a percent of agreement between reviewers when the correct response is not known. For criteria where interrater reliability was less than 80 percent, and/or based on a discussion with the team about the differences in understanding of the proposed criteria, the framework was revised (Appendix D). The purposes of the iterative testing were to improve the relevance of the items within the framework to mental health apps targeting a range of mental health conditions, to standardize the language used to describe these items so that they were well-understood by evaluators with varying levels of expertise, and to develop the accompanying training guide to systematically apply the framework.

Search Approach for Mental Health Apps

We used mental health symptoms and diagnostic categories from the DSM-5 to guide our search and selection of mental health apps. We cross-checked the main diagnostic categories with mental health conditions addressed by current mental health apps and included those addressed by at least one app. In addition to the DSM-5 categories, we added categories for Self-harm; Mental wellness; and Other mental disorders (Table 1).

Table 1. Categorization of mental health disorders

Category*	Included Disorders/Description
Anxiety Disorders	Agoraphobia, anxiety, social anxiety - phobia, separation anxiety disorder, generalized anxiety disorder, panic disorder, panic attack
Bipolar, Depressive Disorder, and Related Disorders	Bipolar disorder, major depressive disorder, mood disorder, depression, dysthymia
Mental Wellness*	Meditation, mindfulness, stress management
Neurocognitive Disorders	Dementia, Alzheimer's
Neurodevelopmental Disorders	Autism Spectrum Disorder, language disorders, learning disorders, traumatic brain injuries, neurogenetic disorders, motor disorders, intellectual disabilities, communication disorders, attention deficit hyperactivity disorder
Obsessive-Compulsive and Related Disorders	Obsessive compulsive disorder, body dysmorphic disorder
Personality Disorders	Borderline personality disorder, personality disorder
Schizophrenia Spectrum and Psychotic Disorders	Schizophrenia, psychosis
Self-Harm*	Suicidal ideation and suicide attempts
Sleep-Wake Disorders	Insomnia, narcolepsy
Substance-Related and Addictive Disorders	Smoking, binge drinking, substance use disorder
Transdiagnostic*	Symptoms across multiple disorders
Trauma- and Stressor-Related Disorders	Post-traumatic stress disorder, adjustment disorders, traumatic stress exposure
Feeding and Eating Disorders	Bulimia, anorexia nervosa, binge-eating disorder, eating disorder
Other mental disorders	Other mental disorders that may not be listed above

*Categories added by Johns Hopkins University Evidence-based Practice Center (i.e., not in DSM)

Apps were identified using 42matters (available at <https://42matters.com/>). 42matters is a database with names, developers, descriptions, and a range of technical details of apps contained in the Apple App Store and Google Play Store. Search criteria on 42matters were limited to apps that fell in the Interactive Advertising Bureau (IAB) categories of “Healthy Living” and “Medical Health”; were released after January 1, 2010; were in English; and were available in the United States.

Search terms used were autism, language disorders, learning disorders, traumatic brain injuries, neurogenetic disorders, motor disorders, intellectual disabilities, communication disorders, attention-deficit/hyperactivity disorder (ADHD), schizophrenia, psychosis, bipolar disorder, mood disorder, depression, dysthymia, agoraphobia, anxiety, social anxiety, phobia, separation anxiety disorder, generalized anxiety disorder, panic disorder, panic attack, obsessive-compulsive disorder (OCD), body dysmorphic disorder, posttraumatic stress disorder (PTSD), adjustment disorders, anorexia, bulimia, binge-eating disorder, eating disorder, insomnia, narcolepsy, addiction, smoking, binge drinking, opioid use disorder, substance use disorder, dementia, Alzheimer’s, borderline personality disorder, personality disorder, self-harm, suicide, mental wellness, meditation, mindfulness, stress, mental illness, and psychotherapy. Searching was conducted on May 31, 2021.

We aimed to identify 50 apps from the Apple App Store (iOS) and 50 apps from the Google Play Store (Android) per mental health category for each of the 15 categories. Screening was based on a description of the app provided on 42matters and considered eligible if the app was about either mental health or wellness and included mental health-related therapeutic content.

For each category, the list of apps was deduplicated using Microsoft Excel. Following this, the apps were randomized within the mental health categories, with the goal of identifying six apps from within each mental health category for further evaluation. The apps were then serially downloaded and further assessed for inclusion. If the app was assessed as ineligible, the next app on the list was selected for evaluation until the target goal of six apps per category was reached. Apps were excluded if they could not be downloaded or opened owing to technical issues, they were unrelated to mental health, they were solely for the purpose of telehealth, an access code was required to use them, or access required a payment of greater than \$100.

Iterative Application and Refinement of the Framework

To test and refine the framework, we conducted one pre-pilot round of testing with the core team, followed by six additional pilot rounds of testing by external evaluators. (See Appendix D.) Initially, we planned to assess 90 apps with one reviewer each but determined that dual review would be most informative for framework refinement. A total of 45 apps were evaluated by at least two reviewers.

Ten apps were tested during the pre-pilot round and interrater reliability was assessed for Section 1 during this round. The pre-pilot round was intended to provide the research team with greater insight on the relevance, flow, and clarity of the framework. Following the pre-pilot round, several changes were made and the framework was iteratively applied by external evaluators.

The reviewers were trained to use the framework using the guide (Appendix E) and accompanying explanations (Appendix F). The initial training took 2 hours. Sixty minutes of follow-up training was conducted after the first pilot round of applications. Across the seven rounds, there were 11 reviewers with varying levels of experience ranging from 2 to 30 years in epidemiology, mental health, project management, technology, and public health. Three of the

evaluators were undergraduate public health students, two were graduate public health students, and one was a doctoral student. Five reviewers had a background in mental health and one in technology.

For pilot rounds one to six, following the pre-pilot, we calculated item-level interrater reliability after each pilot round and averaged for each category of questions in Sections 1 to 3 of the framework. We conducted 2-hour meetings after each pilot round to discuss the items where there was disagreement. We clarified and discussed each of the disagreements with reviewers. If the disagreements were not owing to oversight or error, we made changes to either the question or the responses, or we enhanced the explanation to improve clarity. If certain themes were not relevant to certain types of apps, they were considered for omission. We categorized each change made to the framework in one of seven ways: (1) modified the question language for clarity, (2) removed the question, (3) added a question to further capture the concept, (4) added additional guidance notes, (5) added or consolidated response options, (6) added questions that aren't asked elsewhere, and (7) rearranged the question in the framework sequence.

Peer Review and Public Commentary

Experts in mental health, app developers, and individuals representing stakeholder and user communities were invited to provide external peer review of this Technical Brief. AHRQ task order officers and an associate editor also provided comments. We addressed all reviewer comments and revised the framework and Technical Brief, as appropriate. The peer-reviewed draft Technical Brief was posted on the AHRQ website for 4 weeks to elicit public comment. We addressed all public reviewer comments, revising the text as appropriate and documented responses in a disposition of comments report that will be made available 3 months after AHRQ posts the final Technical Brief on the Effective Health Care website.

Findings

This section is organized into the following subsections:

1. Gaps in existing frameworks
2. Results from Key Informant (KI) interviews
3. Testing and refinement of the framework during pre-pilot and pilot rounds
4. Summary of changes made to the framework during testing
5. Overview of the Framework to Assist Stakeholders in Technology Evaluation for Recovery (FASTER) to Mental Health and Wellness

Gaps in Existing Frameworks

Most frameworks had common criteria assessing aspects of credibility, safety, available features, accessibility features for those with disabilities, clinical evidence foundation, interoperability, usability, efficacy, and data privacy and security. Based on the input from KIs and review of frameworks, the following gaps were identified.

Risk and Credibility

As mental health apps gain traction, it is important to understand the risks posed by each app and balance those against the available evidence on the app's clinical safety, efficacy, security and privacy, and institutional and financial stability. While the existing frameworks include questions about privacy, security, risks, and evidence, they did not attempt to assess the risk posed by the app and assign a safety and credibility rating to it.

Vulnerable Populations

Given that the target users of mental health apps might include vulnerable populations and minors, there is a need for additional checks to ensure that apps address the unique needs of such populations. Such features may include things like how and where security and data use procedures are explained whether there are costs associated with the use of the app that may not be obvious to a user, and whether data will be shared. Currently, there is no clear guidance on the type of checks that should be in place to ensure the safety of vulnerable populations.

Accessibility Features

While existing frameworks include accessibility, they do not distinguish between features that are part of the device's accessibility options and specialized features that may have been developed within the app.⁶⁶ Smart phones and tablet devices provide a host of accessibility features and settings, such as text-to-speech, colorblind color scheme, and text and resolution adjustment. Other features, such as screen reader or adaptation of audio/video content with transcription, can be developed by the app developer to improve user experience. We added specific questions to distinguish between the device's accessibility features and those added by the app developer.

Cultural Sensitivity

Increasingly, mental health apps are being developed to address gaps and the unique lived experiences of various populations. Cultural appropriateness can be an important way in which

apps distinguish themselves. This is particularly relevant for marginalized populations. The frameworks we reviewed do not assess whether the app is targeted at specific cultural groups or uses language that is inclusive of certain populations.

Artificial Intelligence

The use of machine learning and artificial intelligence (AI) has been rapidly growing in mental health apps and poses unique challenges to the safe use of apps. AI has huge potential to transform healthcare and enhance efficiencies and innovation in clinical practice. However, there are challenges with informed consent for use, security and transparency, algorithmic fairness and bias, and effectiveness and privacy.⁵² The Food and Drug Administration (FDA) provides guidance, but no regulation, for mobile apps that use AI.^{45, 46} Existing frameworks do not assess whether and how Machine Learning/AI approaches have been used and whether they incorporate user data and feedback to tailor app content and improve the accuracy and validity of care and wellness recommendations.

Mental Health Features and Function

There are many common features provided by mental health apps, such as mindfulness, safety planning, journaling, automated chatbots, gamification, and social and peer group interactions. A user of the framework may want to prioritize an app with specific features that align with best practices and the therapeutic goals of their population of focus. Most app assessment frameworks do not provide a comprehensive list of these features.

Results From Key Informant Interviews

We completed four 1-hour interviews with 12 KIs, each with relevant expertise (Box 1). Box 2 provides the key components of the framework identified during the interviews.

Box 1. Key Informants' expertise, with number of participants

First set of calls (Round 1)
Patient/Caregiver/Advocate (n=1)
Clinicians (n=3)
Primary care
Psychology
Emergency medicine/app developer*
Health System (n=1)
Veterans' Administration System
Payor (n=3)
Second set of calls (Round 2)
Psychiatrists (one developer of an existing framework and the other a representative from the Food and Drug Administration) (n= 2)
Emergency medicine/app developer (n=1)*
Mental health clinicians/app developers (n=2)

*One person participated in two calls

Key Informant Interviews Round One

We scheduled the first set of KI interviews using zoom with a family member/advocate of those with mental disorders, clinicians from various disciplines, and payers. We held the first zoom meeting with an individual representing the patient/caregiver/advocate perspective and the

second zoom meeting for clinicians and payers. Our key takeaways were (1) apps can fill a mental health service gap in areas where there are limited or no mental health professionals, for those who lack behavioral health insurance, and when someone is on a waitlist for care; (2) apps could be an ideal mode of delivery for accessing psychoeducation to increase mental health knowledge; (3) some apps could serve as a bridge to more traditional mental health care for those with barriers such as mental health stigma or mistrust of mental health providers; and (4) certain apps should have a decision alert or clinical monitoring tool to alert the individual, their support network, or their medical provider if there is a need for additional, more intensive mental health or crisis resources. The clinicians mentioned that they were more comfortable recommending mental health apps that use evidence-based practices (e.g., Cognitive Behavioral Therapy) and those with privacy and consumer protection. They also mentioned that individuals with mild-to-moderate mental health concerns and functional ability that is high enough to engage with apps could potentially benefit from mental health apps. The KIs noted that there is an appeal to the potential of using a framework to take an approach to mental health apps that is more rooted in personalized medicine (i.e., assessing whether a specific app will work for a specific individual). However, the KIs were concerned about safety and adverse effects.

Key Informant Interviews Round Two

The initial draft framework was presented to two separate groups of KIs in a third and fourth zoom meeting to elicit feedback. The third KI meeting included two psychiatrists: one the developer of an existing framework and another representing the FDA (Box 1). The KIs mentioned that there is a need to educate patients and clinicians about the benefits and risks of mental health apps. In terms of new framework components, clinical decision support and risk assessment were highlighted, owing to possible unintended iatrogenic impacts of technology in some individuals with mental health conditions as well as the rapidly changing or deteriorating aspects of some mental health symptoms and conditions. Ideally, mental health apps could detect the need for more intensive services and/or provide direct linkage to a crisis text line or hotline.

The fourth KI call included app developers who were also mental health clinicians and researchers. The KIs mentioned that the draft framework should place more emphasis on the scientific evidence to support the use of specific apps. They also noted that app users need to know where their data are going, specifically there needs to be an ability to opt-in versus opt-out of sharing certain data. In terms of the framework, app efficacy and effectiveness are critically important, as is evidence of doing no harm.

The KIs discussed how the framework could be used. They noted at least three different ways providers could use the framework for decision making: providing guidance on a patient's mental health app selection (i.e., responding to a patient who says, "I am considering using app x for my y, what do you think, doctor?"); recommending an app; and prescribing an app. "Recommend" means to informally suggest the use of an app that may not necessarily be coordinated with existing mental health care, while "prescribe" often involves the provider monitoring the use of an app and coordinating the use of an app with existing care. A prescription code could be needed to access an app.⁶⁷ The framework should aim to clarify the difference between recommending an app and prescribing an app, as reflected by the level of knowledge, scrutiny, and oversight by a provider. One barrier to the recommendation of apps by healthcare providers is the view that if they prescribe a mental health app, providers could be held liable for the risk/harm of apps.⁶⁸

We asked KIs about the use of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) diagnostic categories to classify mental health apps. KIs pointed out that currently there are many apps in the mental health space addressing transdiagnostic symptoms that manifest across diagnostic categories. Thus, it would be hard to place many of the existing apps into specific diagnostic categories; however, this may change in the future since the development of mental health apps is a rapidly evolving area.

Box 2. Ideal components/characteristics of a framework: summary from Key Informant interviews

- Data transparency/privacy and availability of a security policy.
- Evidence from at least one high quality, peer-reviewed randomized clinical trial supporting it.
- Publish credentials and experience of development team.
- Look at the mission and vision of the development organization (e.g., private sector, non-profit, government).
- Important to have a commitment to principles of social justice, equity, and inclusion (e.g., reading level, apps in multiple languages or set to a specific educational level, some apps video-based or audio-based for hearing- or vision-impaired individuals, 508 compliance: safe and accessible for people with disabilities).
- Use culturally sensitive language.
- Show evidence of completion of beta testing by users or patients with mental health conditions.
- Describe features of usability and accessibility.
- Evidence of developer responsiveness (i.e., is the app being regularly updated and are bugs being addressed?).
- Show a clear understanding of outcome targets (e.g., does the app clearly identify what changes/outcomes the user may expect?).
- Specify the app target audience clearly (e.g., adults or adolescents): apps should be transparent about which types of patients/populations benefit, in which settings (e.g., inpatient, outpatient) and under what conditions.
- Access to safety/crisis resources (i.e., does the app connect with crisis hot line [cal] or Crisis Text Line?).
- Identify risks in the app and whether they are communicated to a family member, loved ones, or a provider.
- Document if there have been any adverse effects from the app.
- Highlight if the app has been endorsed by trusted mental health professional association or psychiatric association/agency or mental health advocacy group.
- Document whether existing health plans, Employee Health Program/Employee Assistance Program, service systems, settings offer the app to their stakeholders.
- Document if the app uses an evidence-based clinical model/framework.
- Describe if the app allows for personalization and personalized feedback.
- Convey user experience (e.g., star ratings, reviews by users).
- Be transparent about any costs.
- Consider how app will be accessed by users: apps are sometimes available via health plan platforms like myStrength or Recovery Record or programmatically deployed through an Employee Assistance Program or college campus.

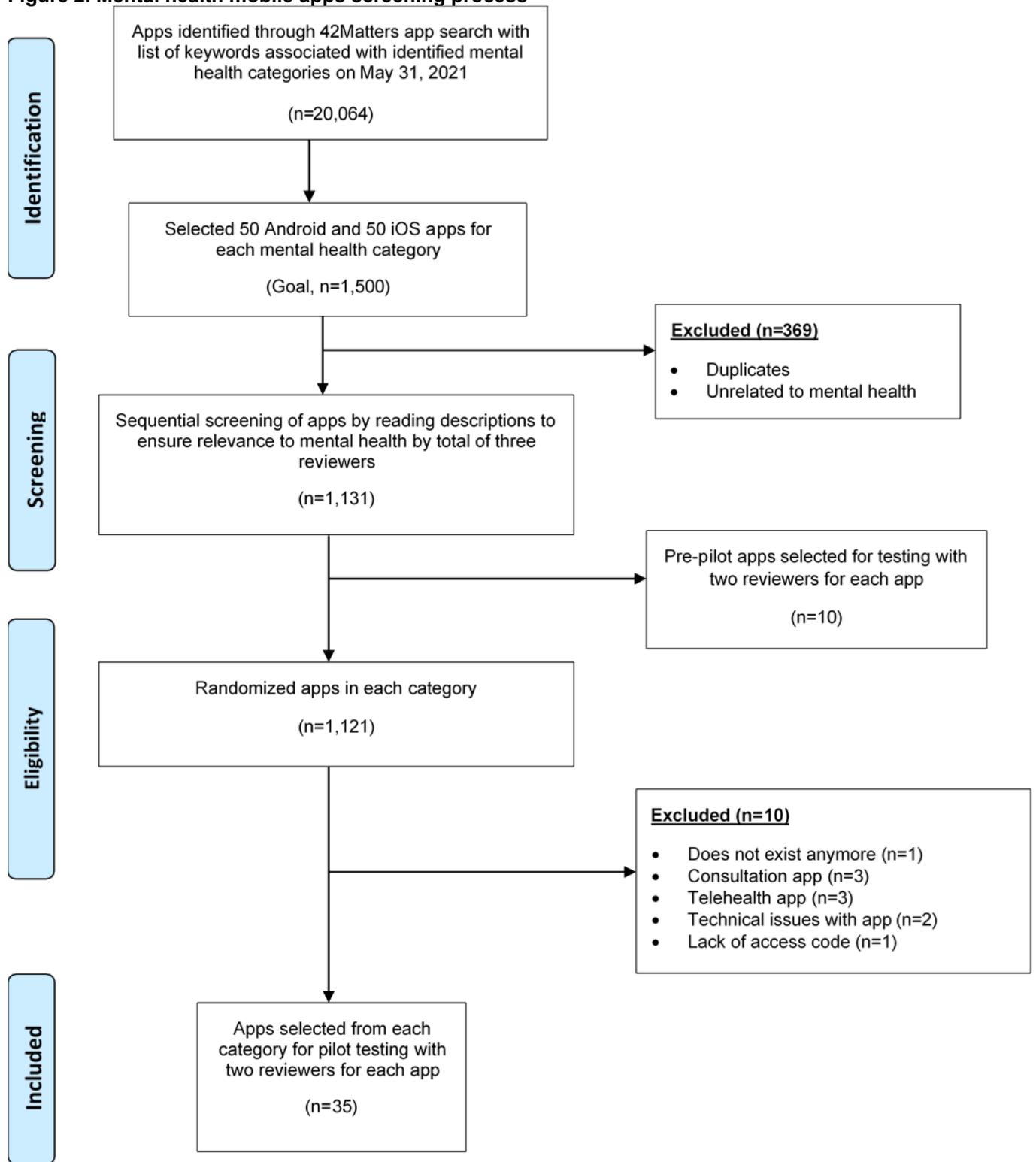
Testing and Refinement of the Framework During Pre-Pilot and Pilot Rounds

Results of the App Search Process on 42matters

We conducted a pre-pilot test of the draft framework using a convenience sample of 10 apps. We selected these 10 apps to represent diversity in (1) the target age group, (2) mental health conditions or symptoms, and (3) app functions. We chose free versions of apps that did not require permission from an employer, healthcare professional, or insurance agency.

Figure 2 describes the results from the app search and screening process for the subsequent testing. Our search yielded 26,064 apps. Our initial aim was to choose 1,500 apps for screening. However, for some mental health categories, we did not reach the target of 50 apps each from Apple and Google store. For example, schizophrenia spectrum and psychosis only had 46 apps from the Google store and 27 from the Apple store, anxiety disorder had 42 apps from the Apple store, eating disorders had 32 from the Apple store and 36 from the Google store, wake-sleep disorder had 37 apps from the Google store, and personality disorders only had 41 apps from the Apple store and 44 apps from the Google store. Additionally, 369 apps were excluded either because they were duplicates or unrelated to mental health. Thus, we screened a total of 1,131 apps. Of these, a convenience sample of 10 apps was assessed in the pre-pilot round. From the remaining 1,121 apps, 10 were excluded as they no longer existed, were telehealth or consultation apps, required an access code, or had technical issues precluding assessment. Thirty-five apps were chosen at random from the mental health categories for the pilot testing. Including the pre-pilot and pilot rounds one to six, we completed dual independent review of 45 apps.

Figure 2. Mental health mobile apps screening process



n = number of apps.

Testing and Refinement

Our core team applied the initial draft framework to 10 apps to acquire a better understanding of the ease of using the framework, the relevance of various criteria, the clarity of the instructions, and the flow of the assessment. After the pre-pilot round, the framework was further iteratively modified in six pilot rounds of testing by evaluators external to the core team. Each app was assessed by two reviewers and interrater reliability was calculated for each question in the framework. The changes made after each round are summarized in Appendix D. We made the most modifications to the Privacy and Security, App Integrity, and Risk Assessment sections. We also rearranged 20 questions between the pre-pilot and pilot round one to create a more sequentially seamless assessment. After pilot round one, we made 48 changes. Changes were made to questions under App Integrity and Privacy and Security sections, with 6 and 8 modifications, respectively. During this round, we had 11 modifications each- 11 required the addition of guidance to clarify the item, and 11 required the addition or consolidation of response options. We also added 9 questions to help us better understand certain concepts. For example, we added “If the app allows for sharing of personal health information (PHI) such as assessments, treatment plan, and messaging with the provider, is it made clear to the user that this data is secured and kept private?” to understand if the use of data is clear to users.

Pilot round two generated half the modifications of pilot round one. Overwhelmingly, our modifications were for adding guidance (8 modifications in pilot round two) or consolidating response options (8 modifications in pilot round two). The total of all other types of modifications was 8 as can be seen in Appendix D, Table D-3. After pilot rounds three and four, we made 5 and 9 modifications respectively, the majority of which involved adding or consolidating response options (3/5 after pilot 3 and 5/9 for pilot 4). The majority of remaining modifications were to add additional guidance notes.

Our final modifications to the framework occurred after pilot round 5. We modified 23 questions, 13 of which were modified by consolidating response options. Many of these questions had “N/A” as an option which created some confusion, so responses were consolidated to allow the framework user to answer the question more easily. Of the remaining modifications, 5 were to add additional guidance notes, 2 questions were rearranged in order, and 3 questions were removed. The Privacy and Security section of the framework was most amended in our final pilot round (10 of 23 modifications), to reflect the difficulty app assessors were having in answering privacy and security questions. After modifications to this section, interrater reliability increased significantly for Privacy and Security questions, from 58 percent to 81 percent between rounds 5 and 6 (Table 2). See examples of assessment in Appendix G.

Table 2. Summary of the interrater reliability in each round and number of changes made to framework to improve interrater reliability

Round	Number of Apps Reviewed	Apps Reviewed	Average Interrater Reliability (%)	Number of Changes to Framework After Each Round
Pre-Pilot	10	Breathe, Think, Do; BUP; Daylio; Replica; Sleepio; SuperBetter; T2 Mood Tracker; The Safe Place; Woebot; Youper	--	106
Pilot Round 1	5	Narcolepsy Monitor; WeAreMore; Awell; Pursue Care; MDLIVE	68.36	48
Pilot Round 2	5	StressScan; Mind Diagnostics; Change Scope; Diet or Disorder; FearTools	62.93	24
Pilot Round 3	4	Self-Harm Recovery; Schizophrenia Health Storylines; Stanley-Brown Safety Plan; PsychSurveys	82.38	5
Pilot Round 4	3	Calm; Joyster; Simple Habit (Sleep, Meditation)	87.51	9
Pilot Round 5	8	Tourette Syndrome; Pattern: Correlate, Health Diary, Mood Tracker; Mental Health Test; Cognitive Stimulation Questions; Quitzilla; Njoy; Narcolepsy Disorder; ShutEye: Sleep Tracker	84.88	23
Pilot Round 6	10	Bipolar Disorder; Mental Health Tracker: Disorders Test; Quirk CBT; Reflectly – Journal and AI Diary; Meditopia; Relaxing Music 2021; Nina App; Breeze: Mood Tracker, Diary; Waking Up: Guided Meditation; PepTalk Motivation	87.03	No changes after round 6

Summary of Changes Made to the Framework During Pilot Testing Rounds

Development of a Pre-Assessment Questionnaire

Several apps that made it past the initial screening for mental health apps were focused entirely on telemedicine or where the primary function was not focused on mental health (e.g., physical fitness tracking apps with only a small component that focused on mindfulness and mental health). The FASTER framework is aimed at evaluating mental health apps. A pre-assessment questionnaire was developed to guide the evaluator on whether the app being evaluated had a mental health focus.

Endorsements and Usage

To assess app integrity, one of the items asked whether the app had been endorsed by a government institution (e.g., Veterans Affairs, Centers for Medicare & Medicaid Services) or a trusted mental health or psychiatric association (e.g., American Psychiatric Association, Substance Abuse and Mental Health Services Administration, National Alliance on Mental Illness). Applying this question to the apps proved to be challenging. Given the legal liability

regarding endorsement, very few apps have endorsements by these types of agencies. There are many apps, such as the Stanley-Brown Safety Plan and Calm, that are used by students and employees at academic institutions and government agencies but not endorsed by any agency. We modified the question to include both endorsement and usage as that captures the intent of the question. By usage we mean that the app may have been made available to employees of the organization. However, we also distinguished that apps do not meet the integrity bar if they are developed by (and not used or endorsed by) governments and academic institutions, given that several apps developed within academic settings are often tested in smaller populations and may not have the funding and institutional backing to scale and be available sustainably to populations.

Assessment of App Risk Based on “Recommended Age”

In the initial rounds of pilot testing, we had lower interrater reliability owing to mis-categorization of risk based on the age groups targeted by the apps. The age at which an app can be used without parental consent varies greatly among apps intended for users 13 to 18 years of age. In the pilot testing, we found that the Apple App Store or Google Play Store rating for recommended age often differed from the permissible age for use listed within the app once it was downloaded. For example, the Calm paid app, which has a Calm for Kids option, states age 4+ in the Apple App Store, yet the Terms of Service say that users over 13 “are not barred from using the Service” and, to purchase the app, users “must be 18 years or older and capable of forming a binding contract.” Further research showed that the App Store rating just indicated the appropriateness of content for a specific age group.⁶⁹ To standardize the evaluation of what age group the apps are targeted towards, we concluded that the Terms of Use, rather than the App Store, is a better source of information for understanding the app’s intended target populations.

“Special” to “Vulnerable” Populations

During testing, we changed the term “special population” to “vulnerable population” and elaborated on our definition of vulnerable populations. We made a distinction between “can the app be used” by these populations to “is the app intended for use” by these populations, highlighting that anyone with access to appropriate technology can use the app, but the intention of this question is to assess whether the app targets a particular group (and if so, does it do so with appropriate checks and balances). Details of the group the app targets are typically listed within the app or its website, and therefore, is more amenable to standard interpretation.

Privacy Policy and Terms of Agreement

There is a lack of standardization of the terminology for warnings, disclaimers, security and privacy agreements, and policies for data sharing used by individual apps. These are distributed, if at all, under “Terms of Use,” “Terms and Conditions,” “Privacy and Security Agreement,” and other agreements that are provided by the app.⁷⁰ The inconsistency in the location of this information required additional guidance for the user of the framework. Another challenge was that sometimes this information was not possible to find in the app itself but could be found on the app website. For the framework, we chose not to include information found on the app website that was not also in the app, because most app users do not go to a website to look up these types of agreements before using the app.

Cultural Competence

Through pilot testing, we broadened the parameters for “cultural competence” to include questions related to groups with certain lived experiences, such as pregnant teens and survivors of gender-based violence. We added clarity around gender inclusivity. We also added a question on whether the app was tested across different ethnic groups and whether efficacy/effectiveness studies of the app included diverse cultural groups.

Changes to Usability Criteria

Of the 20 questions in the Mobile Application Rating Scale (uMARS) framework, a comprehensive set of questions about app engagement and usability, we used 10 questions (Appendix C). To improve clarity, we adapted one of the questions we used: “Quality of information: Is app content correct, well written, and relevant to the goal/topic of the app?” was changed to “Is the app content well written and relevant to the goal and/or topic of the app?”

We observed that the lack of interrater reliability in this category was owing to the differential between the two extremes of response ratings (i.e., evaluators picked either “good” or “very good,” or they chose either “bad” or “very bad” responses; they were typically on the same side of the neutral choice). We made changes to the language used in the responses: in general, we replaced the neutral terms such as “moderate” and “ok” with “satisfactory.” Therefore, for calculation of inter-item agreement, for the categories with five responses, we consolidated the number of responses to three during tabulation by consolidating the two most positive responses and the two most negative responses.

Interaction With a Coach or Healthcare Professional

We added one item based on pilot testing: interaction with a live coach or the ability to engage with a person via text, voice, video, etc.

Categorization of Mental Health Features

Initially, we used a three-point scale for the categorization of mental health features: basic = 1, moderate = 2, and comprehensive = 3. We found that it was difficult to standardize the definitions for each of these labels and reviewers’ agreement on what functionality they considered “basic”, “moderate” and “comprehensive” varied, specifically when differentiating between basic and moderate. To improve interrater reliability, we redefined the features on a two-point scale: not comprehensive = 1, comprehensive = 2.

Acceptable Research

Most apps do not have randomized clinical trials to support their claims of efficacy. Given the rapidly changing and dynamic nature of the field of digital health, we wanted to balance the requirements from a safety perspective while widening the acceptable scientifically validated evidence to evaluate efficacy and effectiveness. Hence, we added an item about research support to determine whether the app included at least one published paper in a peer-reviewed journal that used single-case design or quasi-experimental methods to demonstrate impact.

Subjectivity Versus Objectivity in Question Responses

One of the guiding principles for developing the framework was to restrict the questions and response categories to options that can be objectively assessed. Objective information, while observable, quantifiable, and provable, may not convey quality to the same extent as subjective responses that provide a more nuanced evaluation, even considering the potential for evaluator bias. For example, it is important for the user to know “if the app had advertising that was intrusive and distracting” as well as “if the app had any advertising.” If advertising is a way to support further development of an otherwise free app, it may be acceptable if it isn’t intrusive; having this information allows the user to weigh the pros and cons of the app. We decided to retain some subjectivity for questions with higher interrater reliability but, for those with low interrater reliability, we either consolidated responses or made the response options binary.

Overview of Results of Application of FASTER

As we were evaluating the apps as part of pilot testing and refining the framework, we did not determine consensus ratings or analyze the performance of specific apps. Overall, the reviewers agreed that around 90 percent of the 35 apps reviewed were at Risk Level 1 (low risk) and the rest at Risk Level 2 (higher risk). Only about 10 percent of the apps were assessed by both reviewers as having high integrity. Only a third of the apps provided users with information about resources that could be reached in case of a crisis or an emergency and only one app connected the user automatically to such resources.

We found no apps that had been endorsed by a government or a mental health agency. It was challenging to assess whether an app used an evidence-based strategy, as all apps lacked clear links to published studies. We often found conflicting information around the target patient group for the app. Sometimes apps were described one way on the app store and a different way in their privacy agreement or terms and conditions. Higher quality apps had both a requirement for creating a secure authentication profile (i.e., secure sign-on) and a terms and conditions or privacy statement, though it was usually presented in dense legal language. We found very few apps with an easy-to-follow informed consent process and very few apps that had any features related to cultural competence.

We found a pattern among the poorly developed apps. There was typically no website and hence there was limited information publicly available about them. Often there were no privacy and security agreements on these apps, or the agreements were poorly written and hard to find. Warnings and disclaimers, if at all present, were buried in privacy or terms and conditions contracts. There was also little to no help available to help the user understand how to use the app.

Overview of the FASTER to Mental Health and Wellness Framework

The FASTER to Mental Health and Wellness framework comprises an initial and concluding set of administrative questions along with three sections: Section 1: Risks and Mitigation Strategies; Section 2: Function; and Section 3: Mental Health App Features. (See Appendixes E and F.) Within each of these sections, there are a series of questions related to the assessment of specific categories that were considered critical based upon the literature review and KI interviews. The questions are organized in a systematic order so that a reviewer can start the assessment as they search for and download an app. The rationale for the criteria within each of

these sections is provided in Figure 3, and the categories of questions in each section are shown in Figure 4. Section 1: Risks and Mitigation Strategies facilitates an assessment of the risk profile and integrity of the apps and serves to flag any apps that do not meet basic safety, evidence, and security checks. An app that is flagged for failing Section 1: Risks and Mitigation Strategies may not need to be further assessed.

Figure 3. Framework to Assist Stakeholders in Technology Evaluation for Recovery (FASTER) to Mental Health and H

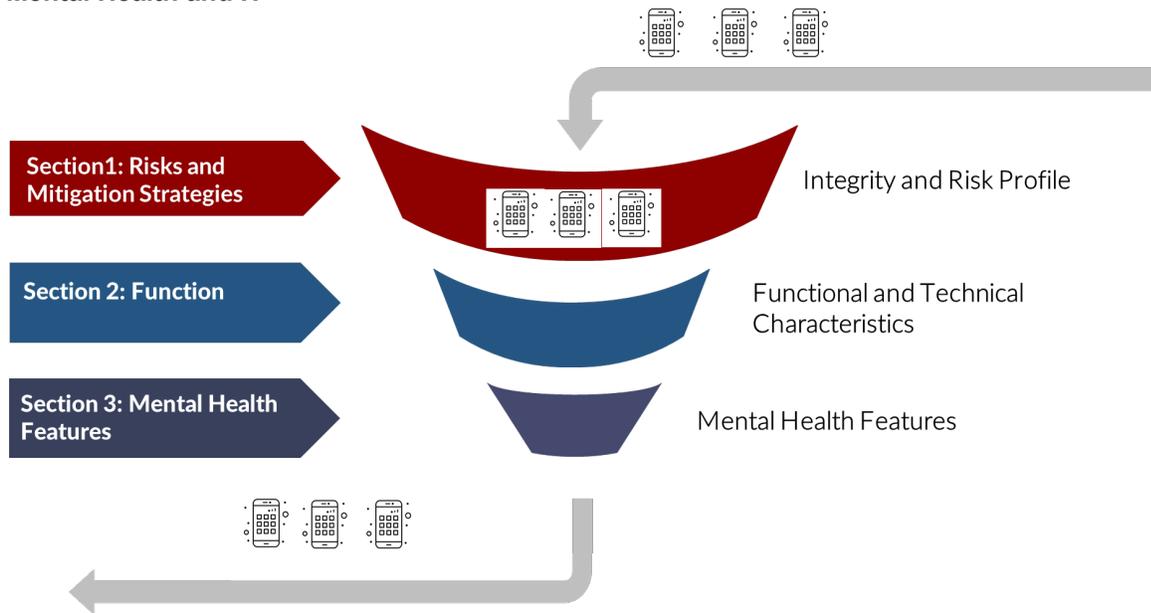


Figure 4. Categories of questions in each section

	Risks and Mitigation Strategies	App Integrity	Risk Assessment	Evidence	Linkage to Care	Access to Crisis Resources
	Function	Accessibility Features	App Info	Costs	Organizational Credibility	Evidence & Clinical Foundation
		Privacy & Security	Informed Consent	Cultural Competence	Usability	Remote Monitoring
		Access to Crisis Resources	Artificial Intelligence (AI)			
	Mental Health Features	Mental Health Features (e.g., facilitating social interaction, motivation enhancement, planning/alternative strategies/planning for high-risk situations, screening, self-help, skill building, safety planning, and promoting sleep hygiene)				

Target Audience and Intended Use

The goal of this assessment framework is to support agencies and individuals serving those with wellness and mental health needs, as well as users of mental health apps, in making

informed decisions about using or recommending the use of particular apps. We expect that this framework will be applied by intermediary mental health advocate groups and entities that have the capacity to train personnel to use this framework to evaluate mental health apps, as well as employers or insurers that might have an interest in coverage or reimbursement of certain apps. Evaluators who use this framework and accompanying training guide should be individuals with some background in technology and mental health, but it is not required that they be experts in these areas. Ideally, the results and summary conclusions of such app assessments using FASTER will be valuable to providers and health systems before they recommend or prescribe apps, and to consumers, family members and peer supports in search of mental health and wellness apps. The framework might also inform and guide app developers in the development of apps.

The FASTER framework is intended to be applied to assess apps whose primary function is to support mental health and wellness through content and resources within the app. It is not appropriate to use this framework to evaluate apps whose primary function is to facilitate telemedicine (e.g., links users to a mental health professional), or apps that might contain cursory content to support wellness (e.g., a weight loss app that has resources for mindfulness).

Administrative Questionnaire

The questions in the administrative section precede the formal evaluation of the app and include introductory questions related to the app that is being reviewed: information about the developer, the website, the country of origin, device compatibility (iOS, Android), approval by any regulatory authorities, the app version number, costs, and requirement of authorization for use. These criteria are factual rather than evaluative and aim to gather data on accessibility, currency, and credibility associated with the use of the app (Table 3). Such information can be important in cataloging and classifying mental health apps as many apps are not static and are being continuously enhanced and updated.

Table 3. Administrative questionnaire

Question	Response
1. Reviewer Name	Free text
2. Date of Evaluation	Free text
3. App Name	Free text
4. App/Company Website	Free text
5. Country of Origin	Free text
6. App Version Number	Free text
7. Platform used for assessment by the reviewer	Free text
8. OS Version Number on Evaluating Device	Free text
9. Describe the goal of the app in your own words. Do not use any of the marketing phrases/lingo used by the app developer.	Free text
10. Has the app been approved by any regulating authority, such as FDA?	Free text
11. Does the app have multiple revenue models, such as freemium, in-app purchases, etc.?	Free text
12. If the app has multiple versions (e.g., free, freemium, paid), which version did you evaluate? If the app had only one revenue model, please choose "not applicable" here.	<ul style="list-style-type: none"> • Free • Paid version • Paid by insurance agency, employer, or healthcare provider • Not applicable • Unable to assess
13. Does the app require prior authorization from a healthcare institution/insurance provider/college/employer or other institution for access?	<ul style="list-style-type: none"> • Yes • No • Unable to assess

Question	Response
14. If you were unable to assess the app, please specify the reason.	Free text

Section 1: Risks and Mitigation Strategies

This section aims to assess the integrity and risk profile of the app. This section is evaluative, and apps that do not meet the thresholds for risks and integrity are flagged. Tables 4 and 5a, 5b, and 5c detail the specific questions within each item. This section contains only those questions that are used for assessment of the risk and app integrity.

Table 4. Section 1: Risk and Mitigation Strategies

Category	Question	Response
App Integrity	1. Was the version of the app you are reviewing updated in the last 6 months?	<ul style="list-style-type: none"> • Yes • No • Unable to assess
	2. Does the app provide a privacy and security agreement to be reviewed and agreed to by the user?	<ul style="list-style-type: none"> • Yes, there is a privacy and security agreement to be agreed to by the user • No, there is not a privacy and security agreement • Unable to assess or unable to access
	3. Does the app provide warnings and disclaimers (e.g., the limitations of the app related to medical liability)?	<ul style="list-style-type: none"> • Yes, warnings and/or disclaimers are provided • No, warnings and/or disclaimers are not provided • Unable to assess • Not applicable
	4. Has the app been endorsed by, or is it being used by, a government agency or trusted mental health professional association?	<ul style="list-style-type: none"> • Yes, the app has been endorsed by one or more mental health associations, government agencies, or non-government bodies • No, the app has not been endorsed • Unable to assess, please specify reason:
App Integrity Assessment	Does the app satisfy the integrity assessment?	<ul style="list-style-type: none"> • Responses to questions 1-3 are all <u>Yes</u> OR response to question 4 is YES: App integrity is "High" • Response to any of the questions 1-3 is No AND response to question 4 is No: App integrity is "Low"
Risk Assessment	5. Is the app intended for use by adults who may have an illness or disability that impacts their decision making ability? Severe autism or severe dementia are examples of illnesses or disabilities that may affect decision-making ability.	<ul style="list-style-type: none"> • Yes • No • Unable to assess
	6. Does the app claim that it is intended for use by minors (i.e., people younger than 18 years of age)?	<ul style="list-style-type: none"> • Yes, for use by minors • Yes, for use by both minors and adults • No, not for use by minors • Unable to assess
	7. Does the app claim to provide standalone treatment for any mental health condition?	<ul style="list-style-type: none"> • Yes, the app provides standalone treatment • No, the app does not provide standalone treatment • Unable to assess

Category	Question	Response
Evidence	8. Has the app been evaluated for efficacy/effectiveness through a scientifically validated study?	<ul style="list-style-type: none"> • Yes, there is strong research support for the app (i.e., at least one published paper in a peer-reviewed journal that uses a randomized trial that shows efficacy or effectiveness) • Yes, there is some research support for the app (i.e., at least one published paper in a peer-reviewed journal that uses single-case design, quasi-experimental methods demonstrating efficacy) • No, there is no research support for the app
	9. Is the app based on or does it use an evidence-based strategy, such as Cognitive Behavioral Therapy (CBT), Dialectical Behavior Therapy (DBT), or evidence-based guidelines?	<ul style="list-style-type: none"> • Yes, the app reports using an evidence-based strategy/guidelines to achieve its goals • No, the app does not report the use of evidence-based strategy/guidelines to achieve its goals (or there are no goals described) • Unable to assess
Linkage to Care	10. Does the app facilitate remote monitoring of the patient or send alerts to a clinician/clinical care team or caregiver?	<ul style="list-style-type: none"> • Yes • No • Unable to assess
	If Yes, please specify who can monitor the health of the patients, either through alerts or by other means	<ul style="list-style-type: none"> • Healthcare provider only • Caregiver only • Both healthcare provider and caregiver • Other • Not applicable
Access to Crisis Resources	11. Does the app provide users with information about resources that can be reached in case of an emergency? This information should be available on the mobile app. Reviewers should not look for this information on the website.	<ul style="list-style-type: none"> • Yes (e.g., a crisis hotline, 911, nearest emergency room services, etc.) • No, the app provides no information on emergency services • Unable to assess
	12. If the app can be used by individuals that require substituted consent OR by minors, then is consent sought from either a caregiver/parent/ legal guardian?	<ul style="list-style-type: none"> • Yes • No • Unable to assess • Not applicable
Risk Assessment	What is the risk category for the app?	<ul style="list-style-type: none"> • Risk Level is 3: Response to EITHER Questions 5 or Question 6 is Yes AND response to Question 7 is Yes, go to Safety Assessment Step 1 • Risk Level is 2: Response to EITHER Questions 5 or Question 6 is Yes OR response to Question 7 is Yes, go to Safety Assessment Step 1 • Risk Level is 1: None of the above, go to Safety Assessment Step 3

Table 5a. Step 1: Preliminary Safety Assessment for Risk Levels 2 and 3

Risk Level	Question	Safety Assessment
Risk Level 2	If responses to Questions 9 and 11 are Yes and Question 10 is Yes or No, supporting evidence available, then:	<ul style="list-style-type: none"> • Prelim security check passed, go to Safety Assessment Step 2
	Else	<ul style="list-style-type: none"> • Prelim security check failed, go to Safety Assessment Step 2
Risk Level 3	If responses to Questions 8-11 are Yes, supporting evidence-base available, then:	<ul style="list-style-type: none"> • Prelim security check passed, go to Safety Assessment Step 2
	Else	<ul style="list-style-type: none"> • Prelim security check failed, go to Safety Assessment Step 2

Table 5b. Step 2: Safety Assessment for Vulnerable Population

Question	Safety Assessment
If response to Question 12 is Yes, then:	Safe for vulnerable populations, go to Safety Assessment Step 3
If response to Question 12 is No, then:	Not safe for vulnerable populations, go to Safety Assessment Step 3
Else	Go to Safety Assessment Step 3

Table 5c. Step 3: Safety Assessment

Risk Level	Prelim Safety Check	Vulnerable Population	App Safety
Risk Level 1	N/A	N/A or SAFE for Vulnerable Populations	Passed Safety Check
	N/A	Not Safe for Vulnerable Populations	Failed Safety Check
Risk Level 2	PASS	N/A or SAFE for Vulnerable Populations	Passed Safety Check
		Not Safe for Vulnerable Populations	Failed Safety Check
	FAIL	--	Failed Safety Check
Risk Level 3	PASS	N/A or SAFE for Vulnerable Populations	Passed Safety Check
		Not Safe for Vulnerable Populations	Failed Safety Check
	FAIL	--	Failed Safety Check

App Integrity

Questions in the App Integrity category assess whether the app can be trusted for use. This is based on whether the app is being updated regularly, based on user feedback and industry trends; whether it provides warnings and disclaimers about use; and whether it gets users to sign a privacy and security agreement if it collects user data. These questions aim to ensure that if users are using an app for mental health or wellness, it is not suddenly pulled off the market, that personal health and financial information is secured, and that the app has legal commitment to user privacy and security. The criteria also assess whether an app has been endorsed or is being used by a trusted federal agency (e.g., Veterans Health Administration), or non-government body (e.g., American Psychiatric Association) which would reinforce credibility, as these institutions exercise due diligence before endorsing or making the app available to their members or constituents. Apps are often continuously evolving through updates. While these updates are often minor, fixing bugs or minor issues related to usability, they are also occasionally substantive, with major changes in functionality. This has implications for an assessment framework in terms of periodically refreshing evaluation of apps. The responses to questions about these concepts determine the level of integrity as High or Low, as defined in Table 6.

Table 6. App integrity levels

Integrity Level	Requirements
High	The app has been updated in the previous 6 months, ensures privacy and security of the user's data (or/and provides disclaimers and warnings), and/or the app has been endorsed by a trusted organization.
Low	The app has not been updated in the previous 6 months and/or provides no privacy and security statement, and/or provides no disclaimers and warnings.

Risk Assessment

Questions in the Risk Assessment category assess the risks posed by a mental health app, as evaluated by alignment of the goals of the app, the target audience, and the severity of the mental health condition, with available evidence to support the approach, oversight and linkage to care, and privacy and security protocols. Linkage to care is defined as linkage to a healthcare provider who can monitor the patient through an interface in the app or through data being linked to

theirelectronic medical record (EMR) system. Additional risk criteria are defined for vulnerable populations.

We adapted alignment between intended goals, target audience, mental health condition, and evidence/safety measures for mental health apps from the model proposed by International Medical Device Regulators Forum “Software as a Medical Device (SaMD): Clinical Evaluation.”⁴⁵ Assessment of the level of risk posed by the app is based on the primary function of the app (standalone treatment vs. other uses, such as psychoeducation and wellness) and whether the app is intended to treat those living with specific mental health diagnoses and/or mental health impairments that substantially impact their ability to function. For example, if the app is used by someone with a more severe clinically diagnosed condition, such as schizophrenia, then the risk posed by the app is higher than if the person has mild self-diagnosed anxiety. Similarly, if the app targets children and adolescents, then the potential risk associated with the app may be higher than if it is an app for adults in the general population. If the app is used by someone with a mild or moderate level of functional impairment, then the risk is greater than an app that can be used by any member of the general population with no functional impairment. Similarly, if an app claims to provide treatment without the support of a healthcare provider, then the risks posed may be greater than those posed by a mindfulness meditation app. Based on these criteria, apps may be classified under three risk levels (Table 7), each of which have specific requirements for evidence and linkage to caregivers or clinical care.

Table 7. Risk levels

Risk Level	Evidence Requirement
Risk Level 1: Minimal Risk	No requirement for providing evidence or for linkage to care. For example, apps aimed at supporting mindfulness practices would fall into this category.
Risk Level 2: Some Risk	Requires some research support regardless of the experimental design. The app should also leverage an evidence-informed theory to guide its approach. Additionally, it should facilitate remote sharing of information with a provider and provide the user with information on a crisis hotline or other resources. For vulnerable populations, the app should require caregiver permission and facilitate sharing of information with them.
Risk Level 3 Considerable Risk	Requires research support with at least one high-quality, peer reviewed randomized controlled trial that shows evidence of impact. The app should also leverage an evidence-informed theory to guide its approach. Additionally, it should facilitate remote sharing of information with a provider and provide the user with information to access a crisis hotline or other resources. For vulnerable populations, the app should require caregiver permission and facilitate sharing of information with them.

At the end of this section, the reviewer will be able to determine whether the app clears the required credibility and risk threshold.

Evidence

Questions in the Evidence category help determine whether the app has a solid clinical evidence foundation. The greater the risk of an app, the greater the burden of evidence. For apps that pose a higher level of risk, the framework requires that there are robust studies assessing the efficacy and risks posed by the apps in order for the app to clear the safety screening.

Linkage to Care

Questions in the Linkage to Care category evaluate the linkages to a healthcare provider who can monitor their patient through an interface in the app or through data being linked to their EMR system. If the app poses a higher level of risk, the framework requires that it also provides resources for linkage to care in order for the app to clear the safety screening.

Access to Crisis Services

Questions in the Access to Crisis Services category evaluate whether the app provides access to information about emergency/crisis services such as a hotline or crisis text line.

Section 2: Function

This section is focused on descriptive aspects related to accessibility, costs, organizational credibility, evidence and clinical foundation, privacy/security, usability, functions for remote monitoring of the user, informed consent, cultural competence, access to crisis services, and AI. These criteria are intended to facilitate (1) systematic cataloging of the functions of the app, so that users may choose an app based on the functionality; (2) a description of features offered by the app so users can assess its fit with their therapeutic and wellness needs; and (3) an assessment of how the app may help individuals with a mental health diagnosis. Questions are outlined in Table 8.

Table 8. Section 2: Function

Category	Question	Response
Accessibility Features	1. Which of the phone's accessibility features work within the app?	<ul style="list-style-type: none"> • Text adjustment feature • Colorblind color scheme feature • Text-to-speech feature • None of the phone's accessibility features work in the app • Unable to assess, please provide a reason:
	2. Are there additional accessibility features that are provided by the app?	<ul style="list-style-type: none"> • Adaptation of audio/video content with transcriptions or captions • Tapping and other gestures are configurable • Contrast text coloring in the content • Screen reader • No additional accessibility features provided by the app • Other features not provided here (please specify): • Unable to assess, please provide a reason:
App Information	3. Does the app work on Apple(iOS)?	Free text
	4. What are the number of reviews on the iOS App Store?	Free text
	5. What is the app rating (number of stars) on Apple Store?	Free text
	6. Does the app work on Android?	Free text
	7. What are the number of reviews on Google Play Store?	Free text
	8. What is the app rating (number of stars) on Google Play Store?	Free text

Category	Question	Response
Costs	9. What is the business model for the app?	<ul style="list-style-type: none"> • Free (no upfront fee for the app; additional packages/services may be offered for a fee) • Up-front fee (a onetime cost for accessing the app; additional packages/services may be offered for additional fees) • In-app purchases (additional packages/services available for a fee in addition to or in lieu of an upfront cost, e.g., concierge services) • Subscription (payment for services on a monthly/quarterly/annual basis) • Reimbursed by healthcare providers/insurers/employers • Other, please specify:
	10. Does the app provide a free or freemium model?	<ul style="list-style-type: none"> • Free • Freemium • No free or freemium version
	11. What is the estimated annual cost of the app for the paid version?	<ul style="list-style-type: none"> • Under \$50 • \$50-250 • Over \$250 • Not applicable
	12. If the app includes paid service(s), does it provide CPT code(s) for insurance reimbursement?	<ul style="list-style-type: none"> • Yes • No or the app has no paid services • Unable to assess, please provide a reason:
Organizational Credibility	13. Who is the developer of the app?	<ul style="list-style-type: none"> • Government • For-profit company • Non-profit institution • Healthcare institution • Academic institution • Insurance company • Independent developer/s • Unable to assess
	14. Does the app have any consumer bureau complaints or lawsuits pending?	<ul style="list-style-type: none"> • Yes • No • Unable to assess
Evidence and Clinical Foundation	15. Does the app appear to do what it claims to do?	<ul style="list-style-type: none"> • Yes, the app provides the functionality it claims on its website • The app provides some of the functionality it claims on its website • No, the app does not provide the functionality it claims on its website • Unable to assess
Privacy and Security	16. Does the app claim it meets HIPAA [or analogous national standard for protected health information (PHI)]?	<ul style="list-style-type: none"> • Yes • No
	17. Does the app claim it meets COPPA [or analogous national standard for protected health information (PHI) for minors younger than 13 years of age]?	<ul style="list-style-type: none"> • Yes • No
	18. Does the app report sharing or selling of data for research or commercial purposes?	<ul style="list-style-type: none"> • Yes • No • Unable to assess

Category	Question	Response
	19. If the app reports sharing or selling for research or commercial purposes, are the data de-identified?	<ul style="list-style-type: none"> • Yes, the app reports that data is de-identified • No, the app does not report data is de-identified • Unable to assess
	20. If the app has the capability to read/write to an electronic health record management system (EHRs) or other healthcare systems, does it use industry standards for secure interoperability (e.g., FHIR, SMART, OAuth 2.0, TLS 1.2)?	<ul style="list-style-type: none"> • Yes, the app uses industry standards for interoperability • No, the app does not use industry standards for interoperability • Not applicable, because it does not read/write to EHRs • Unable to assess
Informed Consent	21. Please assess the level of Consent enabled by the app.	<ul style="list-style-type: none"> • No informed consent • Does not simplify informed consent • Average: Focuses on the essential by a) providing a narrative focused on the most salient information, b) limiting concepts to one per screen, and c) following national plain language and health literacy guidance • Good: Focuses on the essential and organizes content deliberately by a) prioritizing key words/concepts presented on each screen, b) providing information tiers for conceptual elaboration, and c) enabling participants to navigate to their desired level of detail • Excellent: Focuses on the essential, organizes content deliberately, and encourages engagement through interactive elements
	22. What is the format of the data privacy and security consent process followed in the app?	<ul style="list-style-type: none"> • Opt-in for data to be shared, the default choice is opt-out • Opt-out of data sharing, the default choice is opt-in • Hybrid data privacy regime (some opt-in and some opt-out choices) • No choice to opt-out of data sharing, without explicit consent, you can't use the app • No data privacy and security consent process • Unable to assess
Cultural Competence	23. Does the app report developing and testing the app for specific cultural group/s?	<ul style="list-style-type: none"> • Yes • No • Unable to assess • Other, please specify:
	24. If the app reports developing and testing the app for specific cultural groups, is there published documentation (e.g., website, papers) of the process taken to incorporate information that is specific to specified cultures?	<ul style="list-style-type: none"> • Select all that apply: • Website • Documentation (on the app or website) • Published scientific papers • Unable to assess • Other, please specify: • Not applicable
	If the app integrates culturally specific groups, please name the groups.	Free text

Category	Question	Response
	25. Is gender-inclusive language employed? When asking the user's gender, is there an option to self-describe, an option to decline to answer, use of scientifically correct terms for gender (e.g., man, woman, nonbinary)?	<ul style="list-style-type: none"> • Select all that apply: • An option for use of personal pronouns (e.g., they/them, he/his, she/her) • An option to specify another gender • An option to decline to answer • Use of scientifically correct terms for gender, such as man, woman, non-binary • Unable to assess • Not applicable
	26. If the app was tested in a study, what was the percentage of non-white participants?	<ul style="list-style-type: none"> • Less than 30% • Between 30 – 50% • More than 50% • No disaggregate data available • No information available
Usability	27. Does the app work offline?	<ul style="list-style-type: none"> • Yes • No • Unable to assess
	28. Performance: How accurately/fast do the app features (functions) and components (buttons/menus) work?	<ul style="list-style-type: none"> • App is broken; no/insufficient/inaccurate response (e.g., crashes/bugs/broken features) • Some functions work, but app lags or contains major technical problems • App works overall; some technical problems need fixing or app is slow at times • Mostly functional with minor/negligible problems • Perfect/timely response; no technical bugs found or contains a 'loading time left' indicator, if relevant
	29. What languages are supported by the application?	<ul style="list-style-type: none"> • English • Spanish • French • German • Other, please specify:
	30. Customization: Does the app allow customization of settings and preferences by the user (e.g., sound, content, notifications)?	<ul style="list-style-type: none"> • Does not allow any customization or requires setting to be input every time • Allows little customization and that limits the app's functions • Basic customization to function adequately • Allows numerous options for customization • Allows complete tailoring to the user's characteristics/preferences, remembers all settings
	31. Target group: Is the app content (visuals, language, design) appropriate for the target audience?	<ul style="list-style-type: none"> • Completely inappropriate, unclear, or confusing • Mostly inappropriate, unclear, or confusing • Satisfactory, but not specifically designed for the target audience, may be inappropriate/unclear/confusing at times • Designed for the target audience, with minor issues • Designed specifically for the target audience; no issues found
	32. Layout: Is the arrangement and size of buttons, icons, menus, and content on the screen appropriate?	<ul style="list-style-type: none"> • Very bad design: cluttered, some options impossible to select/locate/see/read • Bad design: random, unclear, some options difficult to select/locate/see/read • Satisfactory: few problems with selecting/locating/seeing/reading items • Mostly clear: able to select/locate/see/read items • Professional: simple, clear, orderly, logically organized

Category	Question	Response
	33. Graphics: How high is the quality/resolution of graphics used for buttons, icons, menus, and content?	<ul style="list-style-type: none"> • Graphics appear amateur, very poor visual design: disproportionate, stylistically inconsistent • Low quality/low resolution graphics, low quality visual design: disproportionate • Satisfactory quality graphics and visual design:, generally consistent in style • High quality/resolution graphics and visual design: mostly proportionate, consistent in style • Very high quality/resolution graphics and visual design: proportionate, consistent in style throughout
	34. Visual appeal: How good does the app look?	<ul style="list-style-type: none"> • Ugly/unpleasant to look at: poorly designed, clashing, mismatched colors • Bad: poorly designed, bad use of color, visually boring • Satisfactory: average, neither pleasant nor unpleasant • Pleasant: seamless graphics, consistent and professionally designed • Beautiful: very attractive, memorable, stands out; use of color enhances app features/menus
	35. Does the app have advertising?	<ul style="list-style-type: none"> • Yes • No
	36. If app has advertising, is the advertising intrusive and distracting?	<ul style="list-style-type: none"> • Advertising is neither intrusive nor distracting • The advertising is both intrusive and distracting • Not applicable
	37. Ease of use: How easy is it to learn how to use the app; how clear are the menu labels, icons, and instructions?	<ul style="list-style-type: none"> • No/limited instructions, menu labels and icons are confusing, complicated • Takes a lot of time or effort • Takes some time or effort • Easy to learn (or has clear instructions) • Able to use app immediately, intuitive and simple (no instructions needed)
	38. Navigation: Does moving between screens make sense? Does the app have all necessary links between screens?	<ul style="list-style-type: none"> • No logical connection between screens at all/navigation is difficult • Understandable after a lot of time/effort • Understandable after some time/effort • Easy to understand/navigate • Perfectly logical, easy, clear and intuitive screen flow throughout, and/or has shortcuts
	39. Gestural design: Do taps/swipes/pinches/scrolls make sense? Are they consistent across all components/screens?	<ul style="list-style-type: none"> • Completely inconsistent/confusing • Often inconsistent/confusing • Satisfactory with some inconsistencies/confusing elements • Mostly consistent/intuitive with negligible problems • Perfectly consistent and intuitive
	40. Content: Is the app copy well written and relevant to the goal and/or topic of the app?	<ul style="list-style-type: none"> • There is no information within the app • Irrelevant/inappropriate/incoherent/incorrect • Barely relevant/appropriate/coherent, may be incorrect • Satisfactory with respect to relevance/appropriateness/coherence and appears to be correct • Relevant/appropriate/coherent/correct • Highly relevant, appropriate, coherent, and correct • Not applicable

Category	Question	Response
	41. Is there any evidence to show the average duration of use of the app by users?	<ul style="list-style-type: none"> • On average, users have engaged with the app for 30 days or less • On average, users have engaged with the app for more than a month but less than 6 months • On average, users have engaged with the app for more than 180 days • No evidence available
Remote Monitoring	42. If remote monitoring is feasible, how does the provider access the data?	<ul style="list-style-type: none"> • Provider has access to patient information through the app • Provider has access to the patient information through the EMR • Provider has no access to patient information • Not applicable • Unable to assess • Other, please specify:
	43. Does the app provide alerts to the provider to notify them of a clinical event that may require action on their part?	<ul style="list-style-type: none"> • Yes • No • Unable to assess
	44. Can the app share data with wearables like Apple Watch and Fitbit?	<ul style="list-style-type: none"> • Yes • No • Unable to assess
Access to Crisis Response Services	45. Does the app connect the user automatically to resources in case of a crisis situation or emergency?	<ul style="list-style-type: none"> • Yes • No • Unable to assess
Artificial Intelligence (AI)	46. Does the app use AI?	<ul style="list-style-type: none"> • Yes • No
	47. How is AI being used in the app?	Free text
	48. Is there any evidence to suggest that the app uses data from user interactions to improve precision on AI models?	<ul style="list-style-type: none"> • Yes, there is evidence to suggest that the AI models are updated based on feedback • No, there is no evidence to suggest that the AI models are updated based on feedback • Unable to assess

*Health Insurance Portability and Accountability Act [HIPAA]; Children's Online Privacy Protection Act [COPPA]; Cognitive Behavioral Therapy [CPT]; Fast Healthcare Interoperability Resources [FHIR]; Self-Monitoring Analysis and Reporting Technology [SMART]; Transport Layer Security [TLS]

Accessibility Features

Accessibility features incorporated within an app should facilitate easier use of the technology by individuals with disabilities. Common accessibility features include text-to-speech, closed-captioning, and keyboard shortcuts.

App Information

This section captures details about the platform required by the app (e.g., iOS, Android) and users' perspectives of the apps through the number of reviews and ratings.

Costs

Increasingly, apps have complex pricing models which, especially in the case of a vulnerable user base with mental health impairments, may pose risks. Any costs associated with the app should be provided upfront. The criteria in the Costs category classify the costs as monthly/annual subscription or as freemium services that offer in-app purchasing. Freemium refers to a pricing strategy by which a basic product or service is provided free of charge, but

there are charges for additional features or services that expand the functionality of the free version of the app. Many users of mental health apps may not know that freemium apps may require in-app purchases in order to access the full range of functionality provided by the app. Additionally, some apps require a payment from the user, while others are reimbursed by the healthcare provider or through insurance.

Organizational Credibility

The criteria in the Organizational Credibility category assess any documented complaints against the app developers, as well as the organizational health of the app developing group or company.

Evidence and Clinical Foundation

The questions related to evidence in the Evidence and Clinical Foundation category go beyond what was assessed in Section 1. In this category, the reviewer assesses alignment of the content with the claims made by the app, whether the clinical workflows are rooted in evidence and best-practices, and the clarity of the content. Additionally, the framework assesses evidence about the duration of use of the app because, while some apps are intended to be used for transition periods only, often sustained engagement with apps over time is helpful. This continues to be a significant challenge in the use of apps for healthcare.^{71,72}

Privacy and Security

Security is about the safeguarding of data, including features that provide protection against unauthorized access to data. Privacy is about the safeguarding of the identity of the user. Given the stigma that may be associated with some mental health conditions, privacy of the user is of particular importance for an app that may be developed for vulnerable populations. The criteria to assess privacy and security focus on whether there is transparency about how user data are used for research, quality improvement, or commercial purposes; whether identified or de-identified data are shared or sold to other vendors; how the data are stored; whether user data can be deleted in their entirety (including from group posts); and whether any claims meeting standards of Health Insurance Portability and Accountability Act (HIPAA) or other analogous national standards for PHI are met.

Informed Consent

Informed consent is a process for obtaining permission before conducting some form of research using health data or prior to disclosing the users' health and related information. Most apps have a disclosure list that is long and hard to understand. There are best practices for ensuring that users actually understand what they are agreeing to before they agree to the privacy and security practices. The questions in the Informed Consent category evaluate whether the app follows these best practices.

Cultural Competence

Cultural competence is defined as the ability to understand, appreciate, and account for different cultures or belief systems based on race, ethnicity, income strata, religious beliefs, etc. The criteria in the Cultural Competence category assess whether the app is targeted at, or inclusive of, specific population groups and cultures. If the app is targeted at a specific cultural group, the criteria assess whether the app was tested in that group. The criteria also assess the use of gender inclusive language, and evidence of effectiveness in a non-white population.

Usability

Usability can be described as the capacity of a system to provide a condition for its users to perform tasks safely, effectively, and efficiently. It is important that the user experience be engaging and pleasing, otherwise users are likely to stop using the app. Usability assessments are challenging as they can be fairly subjective. However, objective metrics can measure whether a functionality such as adjustment of font size is present but fails to provide a good assessment of usability. We decided to retain questions that were somewhat subjective for usability.

Functions for Remote Monitoring of the User

Remote patient monitoring is a technology to enable monitoring of patients outside of conventional clinical settings, such as in the home or in a remote area, which could increase access to care and decrease healthcare delivery costs. For mental health apps, the provider may receive an alert about their patient's health, or they may be able to access the patient's health indicators from within the app. To enable remote monitoring, apps need to adhere to established data standards for interoperability to safely exchange health data, including with wearable devices that may be used to monitor vital parameters or behaviors.

Access to Crisis Services

In Section 1, we are assessing if app can place users in contact (24 hours a day, 7 days a week) with a trained counselor from the Crisis Text Line or the National Suicide Prevention Lifeline, 988, or 911 automatically.

Artificial Intelligence

Apps often use or claim to use AI for a variety of tasks including automation, problem solving, and prediction. Apps may claim to use AI when they are just using a rule-based system. In a rule based system the set of rules are coded by a human and result in pre-defined outcomes as opposed to a machine learning system. The questions in the AI category are important because we want to gauge the potential for the apps to cause harm and also determine whether apps are updating algorithms based on user input.

Section 3: Mental Health App Features

This section of the framework aims to match the needs of the user with mental health features certain apps provide. It is a specialized section where the questions may only apply to specific apps. Questions asked include what mental health symptom or condition the app addresses, as well as the primary function of the app (e.g., wellness [mindfulness/meditation/relaxation, psychoeducation]; skills training; symptom tracking/monitoring; social support)(Table 9). Mental health features that apps may include are facilitating social interaction, motivation enhancement, planning/alternative strategies/planning for high-risk situations, screening, self-help, skill building, safety planning, and promoting sleep hygiene. The criteria in this section are intended to facilitate the cataloging of specific features that may enhance or align with therapeutic and wellness goals.

Table 9. Section 3: Mental Health App Features

Category	Question	Response
Mental health app features	1. List mental health symptom(s) and/or condition(s) addressed by the app.	Free text

Category	Question	Response
Mental health app features	2. Please answer the following questions about the app features	<ul style="list-style-type: none"> • Yes • No • Unable to assess
	a. Does the app facilitate text messaging interactions with healthcare therapists, coaches, or other providers via the app?	<ul style="list-style-type: none"> • Yes • No • Unable to assess
	b. Does the app facilitate audio chat interactions with healthcare therapists, coaches, or other providers via the app?	<ul style="list-style-type: none"> • Yes • No • Unable to assess
	c. Does the app facilitate video chat interactions with healthcare therapists, coaches, or other providers via the app?	<ul style="list-style-type: none"> • Yes • No • Unable to assess
	d. Does the app facilitate teletherapy services via the app?	<ul style="list-style-type: none"> • Yes • No • Unable to assess
	e. Does the app facilitate group therapy services via the app?	<ul style="list-style-type: none"> • Yes • No • Unable to assess
	f. Does the app provide live support to a coach or counselor via the app?	<ul style="list-style-type: none"> • Yes • No • Unable to assess
	g. Does the app provide concierge mental health services via the app? Concierge services are personalized services for patients can choose based on what they think works best for them. Some apps may provide acute concierge services to include additional help during times of crisis or high stress.	<ul style="list-style-type: none"> • Yes • No • Unable to assess
	3. Does the app provide a direct connection to 988 or other hotlines?	<ul style="list-style-type: none"> • Yes, and the connection works • Yes, but the connection does not work • No • Unable to assess
Mental health functionality	4. Are the following functionalities supported by the app?	<ul style="list-style-type: none"> • Comprehensive • Not comprehensive • Not applicable
	Mindfulness: Mindfulness is a therapeutic technique that includes elements of relaxation, breathing, and body exercise. It also includes techniques, such as meditation, guided positive imagery, grounding exercises, or progressive muscle relaxation.	<ul style="list-style-type: none"> • Comprehensive • Not comprehensive • Not applicable
	Journaling: Journaling can help users manage anxiety, reduce stress, and cope with mental health challenges. It can provide users with a way of identifying negative thoughts and behaviors and highlight positive aspects in their lives.	<ul style="list-style-type: none"> • Comprehensive • Not comprehensive • Not applicable
	Psychoeducation: Psychoeducation may be defined as the education of a person with a psychiatric disorder regarding the symptoms, treatments, and prognosis of that illness. Psychoeducation may also be targeted to the caregiver, family member, or loved one of the patient. Psychoeducation should be a brief personal intervention by a healthcare provider upon first diagnosis; however, sometimes, only written materials or online resources are provided to the patient. In the worst case scenario, they patients receive no psychoeducation from their provider.	<ul style="list-style-type: none"> • Comprehensive • Not comprehensive • Not applicable

Category	Question	Response
	<p>Skill Building: Skill building includes techniques for recognition of signs and symptoms of a problem, self-advocacy, stress management, emotional regulation, relapse prevention, promotion of sleep hygiene, etc. The app may include tips and advice on dealing with negative emotions and with behavior change. Skill building may require repetition, positive reinforcement, modeling, and practice.</p>	<ul style="list-style-type: none"> • Comprehensive • Not comprehensive • Not applicable
	<p>Screening: Historically, mental health care has relied on structured patient interviews and self-reported questionnaires for diagnosis. These can be used for self-evaluation, reporting, or to decide whether a patient should engage with a mental health professional. Please indicate if industry-validated screening questionnaires (e.g., GATT, PHQ-4, PHQ-9) are provided by the app. The standardized screening questionnaires used will generally be mentioned on the app developer's website.</p>	<ul style="list-style-type: none"> • Comprehensive • Not comprehensive • Not applicable
	<p>Safety Planning: Safety planning is designed to help individuals respond to escalation of suicidal thoughts and feelings, giving them a tailored list of coping behaviors and social support to use until the risk passes.</p>	<ul style="list-style-type: none"> • Comprehensive • Not comprehensive • Not applicable
	<p>Sleep Hygiene: This includes features for promoting sleep hygiene, such as sleep diaries, lifestyle tracking, alarms, data synchronization with a wearable device (e.g., FitBit, Apple Watch), etc.</p>	<ul style="list-style-type: none"> • Comprehensive • Not comprehensive • Not applicable
	<p>Automated Chatbots: Chatbots include conversational agents that can provide virtual therapy and/or social support.</p>	<ul style="list-style-type: none"> • Comprehensive • Not comprehensive • Not applicable
	<p>Family/Caregiver Support: Caregivers take care of the day-to-day needs of patients. Apps can help caregivers monitor the mental state of the patient, identify the early signs of illness, track relapse and deterioration, and help the patient in access services. The family/caregiver may also supervise treatment and provide emotional support to the patient.</p>	<ul style="list-style-type: none"> • Comprehensive • Not comprehensive • Not applicable
	<p>Social and Peer Group Interaction: Individuals with mental health conditions may find it hard to engage socially and may have a small social network/peer group. Apps may provide peer group interactions and/or community support discussion boards for such individuals.</p>	<ul style="list-style-type: none"> • Comprehensive • Not comprehensive • Not applicable
	<p>Gamification: Gamification is the application of typical elements of game playing (e.g., point scoring, competition with others, rules of play) to other areas of activity. Gamification is used to encourage and reward positive changes.</p>	<ul style="list-style-type: none"> • Comprehensive • Not comprehensive • Not applicable
	<p>Personalization: Personalized treatment entails the selection of a therapy or treatment protocol based on a patient's profile, which may increase the likelihood of a successful outcome. Personalization can be used to tailor interactions initially, when the user starts to use the app, and also on a more regular basis. It is usually based on user information collected by the app, such as what activities produce positive emotions (e.g., physical exercise, talking to a friend).</p>	<ul style="list-style-type: none"> • Comprehensive • Not comprehensive • Not applicable

Category	Question	Response
	Self-Monitoring: Self-monitoring apps allow users to engage in symptom monitoring, mood tracking, and progress tracking through a treatment program, such as one for substance abuse, depression and anxiety, etc.	<ul style="list-style-type: none"> • Comprehensive • Not comprehensive • Not applicable
	Medication Adherence: Does the app support patients in their use of medication(s)? This could be through alerts, notifications, or other means.	<ul style="list-style-type: none"> • Comprehensive • Not comprehensive • Not applicable
	Medication Delivery: Does the app support physical delivery of medications to patient(s)?	<ul style="list-style-type: none"> • Comprehensive • Not comprehensive • Not applicable
	5. Please add any features supported by the app that are not listed above.	Free text

Post-Administrative Questionnaire

The Post-Administrative Questionnaire solicits details about the app and any training that is available and required for using the app. Links to all the evidence analyzed as part of the review should be documented here. This questionnaire also affords the reviewer an opportunity to provide a subjective evaluation of the app (Table 10).

Table 10. Post-Administrative Questionnaire

Question	Response
1. Include all links to references for evidence (citations on PubMed, systematic reviews, websites, etc.) here.	Free text
2. Based on your review of the app, do you think it would have been useful to have some training or a tutorial about how to use it?	<ul style="list-style-type: none"> • Yes, there was training available, and I needed it to use the app • Yes, there was training available, but I didn't need it to use the app • No, there was no training available, but I needed it to use the app • No, there was no training available, and I didn't need it to use the app • Don't know
3. Does the app have any help-related documentation available in the app itself (e.g., tooltips, general help)?	<ul style="list-style-type: none"> • Yes, there was comprehensive help available on the app • Yes, there was some help available on the app • The app help button took me to the website • No help was available on the app
4. Does the app have a dedicated website that provides information about the app?	<ul style="list-style-type: none"> • Yes • No • Unable to assess
5. Does the app have any help-related documentation available on its website?	<ul style="list-style-type: none"> • Yes, there was comprehensive help available • Yes, there was some help available • No, help was not available on the app
6. Please provide a subjective evaluation of the app in your own words.	<ul style="list-style-type: none"> • Do you think the assessment conducted here matches your subjective assessment? If not, please provide additional details. • Do you think the risk posed by the app is captured accurately here? • Do you think the app is technically sound? • Do you think the therapeutic content provided by the app is rudimentary or substantive? • Can this app help those that may be experiencing mental or behavioral health challenges?
7. Do you think the app could cause harm or have a negative impact on the user? Please elaborate.	Free text

Question	Response
8. Do you think the subjective risk assessment matches the risk assessment calculated in Section 1? Please elaborate.	Free text
9. Any additional comments?	Free text
10. Please note the time taken to complete your assessment using this framework.	Free text

Summary and Implications

Mental health mobile apps can fill a major gap in mental health services. At present, the evidence base on the efficacy of the various approaches that mental health apps use is limited and there is some potential risks associated with many mental health apps. This field is rapidly evolving, and the COVID-19 pandemic has illuminated the critically important role technology can play to support wellbeing and mental health.

Our proposed framework, Framework to Assist Stakeholders in Technology Evaluation for Recovery (FASTER) to Mental Health and Wellness, is aimed at facilitating the use of apps for mental health support and recovery through standardized evaluation, screening, and classification of apps. We extracted criteria from existing frameworks.^{32-34 27-31, 35,36, 37} However, we identified several gaps in the existing frameworks and addressed these through further prioritization of criteria, the addition of criteria to assess risks and safety of apps, and assessment of the use of artificial intelligence (AI) and other engagement approaches. Prior frameworks to evaluate health apps have been largely focused on the evaluation of the technology specifications rather than the use of the apps in specific healthcare domains. In contrast, FASTER to Mental Health and Wellness is structured to facilitate an initial screening of apps to align the purpose of, and the mental health condition targeted by, an app with the possible risk and risk mitigation features (e.g., foundational evidence, access to crisis hotlines) to determine overall safety. Additionally, FASTER to Mental Health and Wellness includes some guidance to address the vast and rapidly expanding use of AI in apps targeted at mental health.

Next Steps

We envision a range of possible applications of this framework. First, the framework can be used by organizations and advocacy agencies that provide mental health resources to develop a curated and validated app library. Such libraries could be consumer-facing with recommendations tailored to the mental health condition or the type of skill/resource for which the user is seeking assistance. The libraries could also be used by individuals, family members, peer supports, and health care providers to select apps. Updating and maintaining such a library would necessitate substantial planning and resources for routine curation and screening of apps using the framework. Second, the framework can be leveraged by health plans; health system leaders; public and private insurance providers; and other entities, such as professional organizations (e.g., American Psychiatric Association), to review apps relevant to their members and provide guidance and advice that is specific to mental health. A federal agency or other trusted mental health organization could host, automate, update, and disseminate the proposed framework. Ideally, the framework should be a living document that evolves with the field. Lastly, the framework can be used by app developers as guidance to promote transparency in communication about the potential benefits, risks, and evidence to support an app.

Assessment of mental health apps using a framework poses some unique challenges that we anticipate will continue to require attention. Assessment of the risks that an app poses to individuals with mental health conditions is challenging. An individual's mental health can change quickly, which changes the potential risks. We accounted for this in the framework by the inclusion of criteria on appropriate linkage to a provider and other caregivers. Many mental health symptoms are transdiagnostic, and typically apps may aim to support a symptom rather than the disease. Section 1: Risks and Mitigation Strategies of the framework assesses risks based on the type of mental health condition and the aim of the app (i.e., standalone treatment vs. supportive/adjunctive care). However, several mental health apps may aim to target symptoms,

such as anxiety or insomnia, which are common across several mental health conditions. Further refinement of the framework may be needed to address applicability for apps that target transdiagnostic symptoms. Additional criteria may be needed to account for potential harm or iatrogenic impacts of an app, based on the severity or other characteristics of specific mental health conditions or which pertain to specific situations, such as a new diagnosis of a mental health condition where acuity may be unclear. Also, as our knowledge of how apps can be effectively used to address mental health conditions grows, this framework will need to be updated to reflect that knowledge. The FASTER framework was not tested on digital therapeutics apps due to the need for special permission/code to access such apps. As the framework continues to evolve, there would be a need to test it on digital therapeutic apps. Lastly, one of the limitations of this framework is that it does not fully assess the nature or quality of the evidence around the efficacy of a mental health app; rather, it only assesses whether such an efficacy study exists. Assessing the strength and quality of the evidence of an intervention would require a systematic review. Such a comprehensive assessment would require substantial financial as well as time resources and are beyond the scope of how this framework might be implemented.

The framework presents considerations that can be generalized for most mental health apps; as such, it does not facilitate a comprehensive assessment of apps for a specific mental health condition. We acknowledge that apps targeted at specific disorders may benefit from an assessment that is quite specific for those types of apps; however, we also recognize that developing such disease area-specific frameworks is time- and resource-intensive and is unlikely to be a practical approach. As a next step to facilitating adaptation and adoption, it would be valuable to gather user (e.g., provider, patient) input on the value of these summative conclusions in guiding decisions about the use of the apps.

We acknowledge that most patients and providers may not be able to review a detailed assessment report on the utility of a certain app and might find a simple “go/no-go” recommendation easier to process. FASTER does not support the development of a nuanced report on the apps and does not provide an explicit recommendation on whether an app should be used. As new governance and regulations for software as a medical device are formulated, the framework could be adapted for such purposes. Similarly, we recognize that there are ongoing developments in our understanding of appropriate prerequisites for apps from a privacy/security perspective, as well as rapid innovation in the digital health and AI space. We expect that this framework will need to be updated routinely to reflect these areas of growth, especially those specific to AI.

In future versions of the framework, it would be important to add greater input from commercial app developers as they can provide insight regarding the app roadmap and challenges in commercializing health apps. Future application, testing, and refinements to this framework may be required to determine its suitability and reliability across multiple mental health conditions, as well as to account for the rapidly expanding applications of AI in mental health apps. Additionally, adaptations to the framework may be needed to account for policy changes regarding the privacy and security of user information in regulatory contexts.

Clear guidance is needed from the Food and Drug Administration (FDA) regarding which mental health apps are and are not considered medical devices, as well as other policies specific to mental health apps. If an app provides a therapeutic intervention, it could be subject to FDA approval, but the term ‘therapeutic intervention’ is vague and can be misinterpreted. It is also important to establish which mental health conditions are considered by the FDA a ‘disease state’; for example, suicidal thoughts or attempts are not a mental health disorder but rather a

severe symptom of many different mental health conditions. Mental health app utilization and dissemination could be slowed in the absence of such guidance.

Ultimately, to facilitate the adoption and sustainability of this framework, it would be necessary to have a system that can train personnel to apply this framework and screen apps. The results would ideally be hosted as an interactive webpage that can be used by consumers, family members and peer supports, and providers and health systems to select apps. Potential stakeholders to provide such support for the framework may include, but are not limited to, health systems, employee assistance programs, colleges/universities, Substance Abuse and Mental Health Services Administration SAMHSA, American Psychiatric Association, American Psychological Association, National Alliance on Mental Illness (NAMI), Mental Health America, mentalhealth.gov, HelpGuide, World Health Organization (WHO), Anxiety and Depression Association of America, Depression and Bipolar Support Alliance, International Obsessive Compulsive Disorder (OCD) Foundation, National Eating Disorders Association, Post Traumatic Stress Disorder (PTSD) Alliance, Schizophrenia and Related Disorders Alliance of America, Treatment Advocacy Center, Active Minds, the Child Mind Institute, American Foundation for Suicide Prevention, and the American Association of Suicidology. To further facilitate appropriate use of mental health apps in clinical and public health contexts, significant education is necessary across the healthcare ecosystem to convey to end-users, including licensed mental health professionals and other clinicians, the potential benefits and risks of such apps. Such training could be provided as continuing education programming. The FASTER to Mental Health and Wellness framework aims to systematically classify and evaluate mental health apps to provide a resource for users to navigate the digital health ecosystem.

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 Soc Care Community*. 2020 Sep;28(5):1366-79. doi: 10.1111/hsc.12968. PMID: 32115797.
3. Mental health: strengthening our response. <https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response>.
4. Mantani A, Kato T, Furukawa TA, et al. Smartphone Cognitive Behavioral Therapy as an Adjunct to Pharmacotherapy for Refractory Depression: Randomized Controlled Trial. *J Med Internet Res*. 2017 Nov 3;19(11):e373. doi: 10.2196/jmir.8602. PMID: 29101095.
5. 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*. 2017 Oct;16(3):287-98. doi: 10.1002/wps.20472. PMID: 28941113.
6. Firth J, Torous J, Nicholas J, et al. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. *J Affect Disord*. 2017 Aug 15;218:15-22. doi: 10.1016/j.jad.2017.04.046. PMID: 28456072.
7. Price M, Yuen EK, Goetter EM, et al. mHealth: a mechanism to deliver more accessible, more effective mental health care. *Clin Psychol Psychother*. 2014 Sep-Oct;21(5):427-36. doi: 10.1002/cpp.1855. PMID: 23918764.
8. Haque SN. Telehealth Beyond COVID-19. *Psychiatr Serv*. 2021 Jan 1;72(1):100-3. doi: 10.1176/appi.ps.202000368. PMID: 32811284.
9. Powell AC, Torous JB, Firth J, et al. Generating value with mental health apps. *BJPsych Open*. 2020 Feb 5;6(2):e16. doi: 10.1192/bjo.2019.98. PMID: 32019619.
10. Bridges KM. Implicit Bias and Racial Disparities in Health Care. https://www.americanbar.org/groups/crsj/publications/human_rights_magazine_home/the-state-of-healthcare-in-the-united-states/racial-disparities-in-health-care/. 2021.
11. Manuel JJ. Racial/Ethnic and Gender Disparities in Health Care Use and Access. *Health Serv Res*. 2018 Jun;53(3):1407-29. doi: 10.1111/1475-6773.12705. PMID: 28480588.
12. Sealy-Jefferson S, Vickers J, Elam A, et al. Racial and Ethnic Health Disparities and the Affordable Care Act: a Status Update. *J Racial Ethn Health Disparities*. 2015 Dec;2(4):583-8. doi: 10.1007/s40615-015-0113-z. PMID: 26668787.
13. Hahm HC, Cook BL, Ault-Brutus A, et al. Intersection of race-ethnicity and gender in depression care: screening, access, and minimally adequate treatment. *Psychiatr Serv*. 2015 Mar 1;66(3):258-64. doi: 10.1176/appi.ps.201400116. PMID: 25727113.
14. Health Resources and Services Administration. Designated health professional shortage areas statistics. Bureau of Health Workforce. U.S. Department of Health and Human Services [Internet], 2019. (https://ersrs.hrsa.gov/ReportServer/?/HGDW_Reports/). 2019.
15. Wang K, Varma DS, Prospero M. A systematic review of the effectiveness of mobile apps for monitoring and management of mental health symptoms or disorders. *J Psychiatr Res*. 2018 Dec;107:73-8. doi: 10.1016/j.jpsychires.2018.10.006. PMID: 30347316.

16. Wang E, Zweig M. A defining moment for digital behavioral health: Four market trends | Rock Health. <https://rockhealth.com/insights/a-defining-moment-for-digital-behavioral-health-four-market-trends/>. 2021.
17. Mercom. Mental Health Apps Raised \$1.2 Billion in 2020. @mercom. <https://mercomcapital.com/mental-health-apps-raised/>. 2021.
18. Notification of Enforcement Discretion for Telehealth Remote Communications During the COVID-19 Nationwide Public Health Emergency. <https://www.hhs.gov/hipaa/for-professionals/special-topics/emergency-preparedness/notification-enforcement-discretion-telehealth/index.html>.
19. Need mental health help? There are apps for that, but picking the right one is tough. <https://www.latimes.com/california/story/2021-06-21/mental-health-apps-consumer-challenge-picking-the-right-one>. 2021.
20. Weir K. The ascent of digital therapies. *Monitor on Psychology*, 49(10). <http://www.apa.org/monitor/2018/11/cover-digital-therapies>. 2018.
21. U.S. Food & Drug Administration. Digital Health Software Precertification (Pre-Cert) Program. <https://www.fda.gov/medical-devices/digital-health-center-excellence/digital-health-software-precertification-pre-cert-program>.
22. Software as a Medical Device (SaMD). FDA. Published September 9, 2020. Accessed May 11, 2021. <https://www.fda.gov/medical-devices/digital-health-center-excellence/software-medical-device-samd>.
23. Bergin A, Davies EB. Technology Matters: Mental health apps - separating the wheat from the chaff. *Child Adolesc Ment Health*. 2020 Feb;25(1):51-3. doi: 10.1111/camh.12363. PMID: 32063750.
24. Martinengo L, Van Galen L, Lum E, et al. Suicide prevention and depression apps' suicide risk assessment and management: a systematic assessment of adherence to clinical guidelines. *BMC Med*. 2019 Dec 19;17(1):231. doi: 10.1186/s12916-019-1461-z. PMID: 31852455.
25. Dehling T, Gao F, Schneider S, et al. Exploring the Far Side of Mobile Health: Information Security and Privacy of Mobile Health Apps on iOS and Android. *JMIR Mhealth Uhealth*. 2015 Jan 19;3(1):e8. doi: 10.2196/mhealth.3672. PMID: 25599627.
26. Akbar S, Coiera, Enrico, Magrabi F. Safety concerns with consumer-facing mobile health applications and their consequences: a scoping review. *Journal of the American Medical Informatics Association*. 2019;27(2):330-40. doi: 10.1093/jamia/ocz175.
27. App Guide | PsyberGuide. One Mind PsyberGuide. Accessed May 11, 2021. <https://onemindpsyberguide.org/apps/>.
28. APA App Advisor. Accessed May 11, 2021. <https://www.psychiatry.org/psychiatrists/practice/mental-health-apps>.
29. How to pick employee mental health apps | Kaiser Permanente. Accessed May 11, 2021. <https://business.kaiserpermanente.org/insights/mental-health-workplace/mental-health-apps-workforce-wellness>.
30. Know more. Live brighter. Verywell Mind. Accessed May 11, 2021. <https://www.verywellmind.com/>.
31. New Zealand health information | Health Navigator NZ. Accessed May 11, 2021. <https://www.healthnavigator.org.nz/>.
32. Methodology | myhealthapps.net. Accessed May 11, 2021. <http://myhealthapps.net/methodology>.
33. Healthy Living Apps Guide. Accessed May 11, 2021. <http://www.vichealth.vic.gov.au/media-and-resources/vichealth-apps/healthy-living-apps>.
34. Digital Technology Assessment Criteria (DTAC). NHSX. Accessed May 11, 2021. <https://www.nhsx.nhs.uk/key-tools-and-info/digital-technology-assessment-criteria-dtac/>.
35. Policymakers. Digital Therapeutics Alliance. Accessed May 11, 2021. <http://dtxalliance.org/value-of-dtx/policymakers/>.

36. Lagan S, Sandler L, Torous J. Evaluating evaluation frameworks: a scoping review of frameworks for assessing health apps. *BMJ Open*. 2021 Mar 19;11(3):e047001. doi: 10.1136/bmjopen-2020-047001. PMID: 33741674.
37. Stoyanov SR, Hides L, Kavanagh DJ, et al. Development and Validation of the User Version of the Mobile Application Rating Scale (uMARS). *JMIR Mhealth Uhealth*. 2016 Jun 10;4(2):e72. doi: 10.2196/mhealth.5849. PMID: 27287964.
38. BfArM - Germany Federal Institute for Drugs and Medical Devices. <https://www.emergobyul.com/resources/urope/germany-federal-institute-drugs-medical-devices>.
39. NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE EVIDENCE STANDARDS FRAMEWORK FOR DIGITAL HEALTH TECHNOLOGIES. <https://www.nice.org.uk/Media/Default/About/what-we-do/our-programmes/evidence-standards-framework/digital-evidence-standards-framework.pdf>. 2019.
40. Celik H, Lagro-Janssen TA, Widdershoven GG, et al. Bringing gender sensitivity into healthcare practice: a systematic review. *Patient Educ Couns*. 2011 Aug;84(2):143-9. doi: 10.1016/j.pec.2010.07.016. PMID: 20719461.
41. Ferreira S, Pêgo JM, Morgado P. The efficacy of biofeedback approaches for obsessive-compulsive and related disorders: A systematic review and meta-analysis. *Psychiatry Res*. 2019 Feb;272:237-45. doi: 10.1016/j.psychres.2018.12.096. PMID: 30590278.
42. Sander LB, Schorndanner J, Terhorst Y, et al. 'Help for trauma from the app stores?' A systematic review and standardised rating of apps for Post-Traumatic Stress Disorder (PTSD). *Eur J Psychotraumatol*. 2020;11(1):1701788. doi: 10.1080/20008198.2019.1701788. PMID: 32002136.
43. Guo Y, Yang F, Hu F, et al. Existing Mobile Phone Apps for Self-Care Management of People With Alzheimer Disease and Related Dementias: Systematic Analysis. *JMIR Aging*. 2020 Jan 24;3(1):e15290. doi: 10.2196/15290. PMID: 32012045.
44. Webb TL, Joseph J, Yardley L, et al. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *J Med Internet Res*. 2010 Feb 17;12(1):e4. doi: 10.2196/jmir.1376. PMID: 20164043.
45. Software as a Medical Device (SaMD): Clinical Evaluation. Accessed June 8, 2021. <http://www.imdrf.org/consultations/cons-samd-ce.asp>.
46. Health C for D and R. Device Software Functions Including Mobile Medical Applications. FDA. Published September 9, 2020. Accessed December 16, 2021. <https://www.fda.gov/medical-devices/digital-health-center-excellence/device-software-functions-including-mobile-medical-applications>.
47. . BfArM - Digital Health Applications (DiGA). Accessed December 12, 2021. https://www.bfarm.de/EN/Medical-devices/Tasks/Digital-Health-Applications/_node.html.
48. Evidence standards framework for digital health technologies | Our programmes | What we do | About. NICE. Accessed May 11, 2021. <https://www.nice.org.uk/about/what-we-do/our-programmes/evidence-standards-framework-for-digital-health-technologies>.
49. Policy for Device Software Functions and Mobile Medical Applications - Guidance for Industry and Food and Drug Administration Staff. :45.
50. Clinical Decision Support Software. U.S. Food and Drug Administration. Published May 6, 2020. Accessed May 11, 2021. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-decision-support-software>.
51. D'Alfonso S. AI in mental health. *Curr Opin Psychol*. 2020 Dec;36:112-7. doi: 10.1016/j.copsyc.2020.04.005. PMID: 32604065.

52. Gerke S, Minssen T, Cohen G. Chapter 12 - Ethical and legal challenges of artificial intelligence-driven healthcare. In: Bohr A, Memarzadeh K, eds. Artificial Intelligence in Healthcare. Academic Press; 2020:295-336.
53. Clinical Decision Support Software. Draft Guidance for Industry and Food and Drug Administration Staff. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-decision-support-software>.
54. IMDRF -International Medical Device Regulators Forum. Software as a Medical Device": Possible Framework for Risk Categorization and Corresponding Considerations. <http://www.imdrf.org/docs/imdrf/final/technical/imdrf-tech-140918-samd-framework-risk-categorization-141013.pdf> 2014.
55. Doerr M, Suver C, Wilbanks J. Developing a Transparent, Participant-Navigated Electronic Informed Consent for Mobile-Mediated Research . Available at SSRN: <https://ssrn.com/abstract=2769129> or <http://dx.doi.org/10.2139/ssrn.2769129>. 2016.
56. Informed Consent for Web Apps. Sage Bionetworks. Accessed December 12, 2021. https://sagebionetworks.org/tools_resources/econsent-app-devs19/.
57. Cultural Competence In Health and Human Services | NPIN. Accessed May 18, 2021. <https://npin.cdc.gov/pages/cultural-competence>.
58. Kahn, J., & Johns Hopkins Project on Ethics and Governance of Digital Contact Tracing Technologies (2020). Digital Contact Tracing for Pandemic Response: Ethics and Governance Guidance. Baltimore: Johns Hopkins University Press., doi:10.1353/book.75831.
59. van Haasteren A, Gille F, Fadda M, et al. Development of the mHealth App Trustworthiness checklist. Digit Health. 2019 Jan-Dec;5:2055207619886463. doi: 10.1177/2055207619886463. PMID: 31803490.
60. SAMHSA -Substance Abuse and Mental Health Services Administration. Resources on Cultural Competency <https://www.samhsa.gov/section-223/cultural-competency/resources>.
61. Substance Abuse and Mental Health Services Administration. Improving Cultural Competence. Treatment Improvement Protocol (TIP) Series No. 59. HHS Publication No. (SMA) 14-4849. Rockville, MD: Substance Abuse and Mental Health Services Administration, 2014. . 2014.
62. Michie S, Richardson M, Johnston M, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med. 2013 Aug;46(1):81-95. doi: 10.1007/s12160-013-9486-6. PMID: 23512568.
63. Riley WT, Rivera DE, Atienza AA, et al. Health behavior models in the age of mobile interventions: are our theories up to the task? Transl Behav Med. 2011 Mar;1(1):53-71. doi: 10.1007/s13142-011-0021-7. PMID: 21796270.
64. Riley WT, Martin CA, Rivera DE, et al. Development of a dynamic computational model of social cognitive theory. Transl Behav Med. 2016 Dec;6(4):483-95. doi: 10.1007/s13142-015-0356-6. PMID: 27848208.
65. Aji M, Gordon C, Peters D, et al. Exploring User Needs and Preferences for Mobile Apps for Sleep Disturbance: Mixed Methods Study. JMIR Ment Health. 2019 May 24;6(5):e13895. doi: 10.2196/13895. PMID: 31127714.
66. AARP -Accessibility Tools on Your Smartphone That Can Make Your Life Easier. <https://www.aarp.org/home-family/personal-technology/info-2020/smartphone-accessibility.html>.
67. Prescribe vs Recommend - What's the difference? <https://wikidiff.com/recommend/prescribe>.

68. Kao CK, Liebovitz DM. Consumer Mobile Health Apps: Current State, Barriers, and Future Directions. *PM R*. 2017 May;9(5S):S106-S15. doi: 10.1016/j.pmrj.2017.02.018. PMID: 28527495.
69. Mobile software content rating system. https://en.wikipedia.org/wiki/Mobile_software_content_rating_system.
70. Robillard JM, Feng TL, Sporn AB, et al. Availability, readability, and content of privacy policies and terms of agreements of mental health apps. *Internet Interventions*. 2019 2019/09/01;17:100243. doi: <https://doi.org/10.1016/j.invent.2019.100243>
71. Melcher J, Camacho E, Lagan S, et al. College student engagement with mental health apps: analysis of barriers to sustained use. *J Am Coll Health*. 2020 Oct 13:1-7. doi: 10.1080/07448481.2020.1825225. PMID: 33048626.
72. Li J, Chang X. Improving mobile health apps usage: a quantitative study on mPower data of Parkinson's disease. *Information Technology & People*. 2021;34(1):399-420. doi: 10.1108/ITP-07-2019-0366.

Appendix A. List of Existing Frameworks

Table A-1. List of existing mental health mobile application assessment frameworks

Framework	Description of Framework	Audience	Conditions	Framework Sections
	Developer/Funder			Number of Questions
One Mind - One Mind Cyber Guide https://onemindpsyberguide.org/apps/	This guide evaluates technology and promotes best practices for the use and development of digital mental health tools with the intention of improving accessibility of mental health resources. Funded by One Mind, an NGO in brain health research (operates out of the University of California, Irvine and Northwestern University)	Employers and end-users	Cognitive Behavioral Principles Psychoeducation/ Symptom Tracking/Self-Assessments/Screening Cognitive Training Dialectical Behavior Therapy Chatbot/AI Goal Setting Connect to a Peer Assessment/Screening Schizophrenia Stress and Anxiety PTSD Mood Disorders Obsessive Compulsive Disorder Phobias Eating Disorders Borderline Personality Disorder Chronic Pain Sleep Substance Use or Addiction Stress & Anxiety	The assessment criteria is focused on 3 core areas. Credibility: combines information about research, development, purpose, and popularity. User Experience: is an app quality score. The Mobile App Rating Scale (MARS) is used to assess this. Transparency: relates to information regarding an apps' data storage and collection policies and how readily available this information is to users 33 Questions

Framework	Description of Framework	Audience	Conditions	Framework Sections
	Developer/Funder			Number of Questions
DTX Product Best Practices https://dtxalliance.org/value-of-dtx/policymakers/	<p>It is a general framework for digital therapeutics that is meant to increase the use of clinically evaluated digital therapeutics with patients, clinicians, payors, policymakers.</p> <p>The Digital Therapeutics Alliance is a 501(c)(6) non-profit trade association that is made up of industry leaders and stakeholders that are involved with creating evidence-based advancement of digital therapeutics.</p>	End-user, clinicians, payors, and policymakers	Digital therapeutics are clinically proven to treat, manage, and prevent a wide range of diseases and disorders using software-based technologies.	<p>The following are the 10 core principles of digital therapeutics.</p> <ul style="list-style-type: none"> Prevent, manage, or treat a medical disorder or disease Produce a medical intervention that is driven by software and delivered via software or complementary hardware, medical device, service, or medication Incorporate design, manufacture, and quality best practices Engage end users in product development and usability processes Incorporate patient privacy and security protections Apply product deployment, management, and maintenance best practices Publish trial results inclusive of clinically meaningful outcomes in peer-reviewed journals Be reviewed and cleared or approved by regulatory bodies as required to support product claims of risk, efficacy, and intended use Make claims appropriate to clinical evaluation and regulatory status Collect, analyze, and apply real world evidence and product performance data <p>10 Questions</p>

Framework	Description of Framework	Audience	Conditions	Framework Sections
	Developer/Funder			Number of Questions
uMARS https://mhealth.jmir.org/2016/2/e72/	<p>This framework is the User Version of the Mobile Application Rating Scale (uMARS) since the framework MARS was not for the end-user.</p> <p>The Young and Well Cooperative Research Centre (Young and Well CRC) funded the project, which is an Australian-based, international research center that unites young people with researchers, practitioners, innovators, and policy-makers from over 70 partner organizations.</p>	End-users	Doesn't specify	<p>The assessment is broken up into 5 subscales.</p> <p>Engagement Entertainment Interest Customization Interactivity Target group</p> <p>Functionality Performance Ease of use Navigation Gestural design</p> <p>Aesthetics Layout Graphics Visual appeal</p> <p>Information Quality of information Quantity of information Visual information Credibility of source</p> <p>Total uMARS Subjective items Would you recommend How many times Would you pay Overall (star) rating</p> <p>20 Questions</p>

Framework	Description of Framework	Audience	Conditions	Framework Sections
	Developer/Funder			Number of Questions
MARS https://mhealth.jmir.org/2015/1/e27/	<p>This framework is the Mobile App Rating Scale (MARS), which is a tool for assessing the quality of health mobile apps</p> <p>The Young and Well Cooperative Research Centre (Young and Well CRC) funded the project, which is an Australian-based, international research center that unites young people with researchers, practitioners, innovators, and policy-makers from over 70 partner organizations.</p>	Researchers, professionals, and clinicians	Doesn't specify	<p>The assessment is broken up into 5 subscales.</p> <p>Aesthetics: graphics, layout, visual appeal</p> <p>Engagement: entertainment, customization, interactivity, fit to target group, etc.</p> <p>Functionality: performance, navigation, gestural design, ease of use</p> <p>Information: quality, quantity, visual information, credibility, goals, description</p> <p>Subjective Quality: worth recommending, stimulates repeat use, overall satisfaction rating</p> <p>Also includes app classification, confidentiality, security, registration, community, affiliation</p> <p>23 Questions</p>

Framework	Description of Framework	Audience	Conditions	Framework Sections
	Developer/Funder			Number of Questions
<p>American Psychiatric Association (APA) App Advisor</p> <p>https://www.psychiatry.org/psychiatrists/practice/mental-health-apps</p>	<p>This framework is meant to help mental health professionals by directing their attention to areas that should be focused on in the selection of mental health apps for their patients. It addresses: the importance of assessing an app, a method for assessing an app, and how to find additional help on apps and/or the evaluation.</p> <p>This model was devised by an Expert Panel that includes American Psychiatric Association-member psychiatric physicians, other mental health professionals (social workers, nurse practitioners), informaticists, medical students, and patients.</p>	<p>Psychiatrists and mental health professionals</p>	<p>Mental health. Doesn't specify conditions.</p>	<p>This brief version of the Model extracts a sample of the most fundamental questions that should be asked before considering using an app.</p> <p>On which platforms/operating systems does the app work? Does it also work on a desktop computer?</p> <p>Has the app been updated in the last 180 days?</p> <p>Is there a transparent privacy policy that is clear and accessible before use?</p> <p>Does the app collect, use, and/or transmit sensitive data? If yes, does it claim to do so securely?</p> <p>Is there evidence of specific benefit from academic institutions, end user feedback, or research studies?</p> <p>Does the app have a clinical/recovery foundation relevant to your intended use?</p> <p>Does the app seem easy to use?</p> <p>Can data be easily shared and interpreted in a way that's consistent with the stated purpose of the app?</p> <p>Brief version: 8 Questions Comprehensive: 37 Questions</p>

Framework	Description of Framework	Audience	Conditions	Framework Sections
	Developer/Funder			Number of Questions
Kaiser Permanente https://business.kaiserpermanente.org/insights/mental-health-workplace/mental-health-apps-workforce-wellness	<p>This is meant as an aid for selecting a mental health app for a workforce.</p> <p>This was determined by a team of Kaiser Permanente health researchers that have been evaluating mental health and wellness apps.</p>	Employers	Meditation and CBT	<p>This guide consists of 5 tips for selecting mental health apps. Narrow your search to meditation and mindfulness and cognitive behavioral therapy apps Test it out before you buy Pick an app that won't share your employees' data Weigh pros and cons of self-guided vs. coaching apps Go to nonprofit and mental health groups for reviews</p> <p>Doesn't ask questions, but gives the above 5 tips</p>
VeryWellMind https://www.verywellmind.com/our-editorial-process-4778006	<p>The purpose of this is to help find appropriate apps that are beneficial to the end-user.</p> <p>This group is by experts in their disciplines, including physicians, therapists, and mindfulness experts.</p>	End-users	<p>Lists features to look for in each type of app listed below. Breathing, meditation, Mindfulness Anxiety Depression COVID-19 Post-traumatic stress disorder Mental health and wellbeing app Goal setting, problem solving and motivation Mental health apps for pregnancy</p>	<p>The evaluation involves two components. Features Clinical Review</p> <p>Doesn't ask specific questions. It examines the features and undergoes clinical review.</p>

Framework	Description of Framework	Audience	Conditions	Framework Sections
	Developer/Funder			Number of Questions
Health Navigator New Zealand: https://www.healthnavigator.org.nz/apps/m/mental-health-and-wellbeing-apps/ Evaluation Methodology: https://www.healthnavigator.org.nz/apps/p/people-process/	<p>This is a general site with information around wellness and health, including apps (not just mental health). It is for finding safe, useful, and relevant health apps.</p> <p>The Health Navigator website is designed by a non-profit community initiative with many partners and supporters, and it is overseen by the Health Navigator Charitable Trust. The app library is supported by the Ministry of Health and is overseen by the New Zealand App Library Advisory Group.</p>	End-users and clinicians	App Selection Process a literature review of research papers published reviews of individual or categories of apps on other independent app review websites searches on Apple and Google Play app stores trending apps on social media and popular news being alerted by app developers about new, relevant apps through our supporter organizations and networks through website users, including consumers using the apps.	<p>Apps designed to meet the following areas more likely to be reviewed</p> <p>Usability Security & privacy General (more details in additional sheet) During the internal review, the following are considered. Features Functionality Information quality</p> <p>Doesn't list questions.</p>
My Health Apps http://myhealthapps.net/ Methodology - http://myhealthapps.net/methodology	<p>This includes general health apps that have been selected as a favorite by end-users and also have app developers that have been transparent about the app.</p> <p>This site does not really explain who the Developer/Funder are.</p>	General public, end-users, carers	ADHD Anxiety Autism-spectrum disorder Bullying Coping with society at large Dealing with a crisis Depression Eating disorders General mental health Obsessive-compulsive disorder Panic disorder Phobia Schizophrenia Sleep Stress Suicide prevention	<p>Two main factors dictate whether a health app is featured on this website. The health app has been nominated as a favorite by patient/disability/carer/family/consumer groups, or by empowered consumers (e.g. consumer advocates, active members/bloggers of moderated consumer health forums) The app developer is transparent about the nature of the app.</p> <p>6 Questions</p>

Framework	Description of Framework	Audience	Conditions	Framework Sections
	Developer/Funder			Number of Questions
Health Living App Guide https://www.vichealth.vic.gov.au/media-and-resources/vichealth-apps/healthy-living-apps	This guide focused on Australians to help find apps that help users change their behaviors to develop a healthier lifestyle. These are health and well-being app focused on the community. It is funded by Victoria Department of Health	Doesn't specify audience, but could be clinicians and end-users	Healthy Living	Behavior change apps were rated for two components. Functionality: rated using the MARS Potential to affect behavior change: rated behavior change effectiveness using the App Behavior Change Scale (ABACUS), developed by McKay and colleagues at Deakin University (McKay et al. 2019). 44 Questions
UK National Health Service https://www.nhs.uk/apps-library/category/mental-health/ App Assessment https://www.nhsx.nhs.uk/key-tools-and-info/designing-and-building-products-and-services/	The Digital Technology Assessment Criteria for health and social care (DTAC) helps clinicians and end-users evaluate whether health apps, not just mental health apps, meet certain standards in terms of clinical safety, data protection, technical security, interoperability and usability and accessibility standards. NHSX is a joint unit of NHS England and the Department of Health and Social Care, supporting local NHS and care organizations.	Clinicians, end-users, and citizens	Mental health/wellness, in addition to general health	The assessment criteria is focused on 5 core areas. Sections 1 to 4 form the assessed criteria, with a separate conformity rating provided around usability and accessibility: Clinical safety: assessed to ensure that baseline clinical safety measures are in place and that organizations undertake clinical risk management activities to manage this risk Data protection: assessed to ensure that data protection and privacy is 'by design' and the rights of individuals are protected Technical assurance: assessed to ensure that products are secure and stable. Interoperability: assessed to ensure that data is communicated accurately and quickly whilst staying safe and secure. Usability and accessibility: products are allocated a conformity rating having been benchmarked against good practice and the NHS service standard. 31 Questions

Framework	Description of Framework	Audience	Conditions	Framework Sections
	Developer/Funder			Number of Questions
MIND (Mhealth Index and Navigation Database) https://bmjopen.bmj.com/content/11/3/e047001	This framework was developed to build off an existing American Psychiatric Association’s (APA) App Evaluation Model framework that their team had developed previously; however, they wanted to create a framework with “more concrete information instead of placing the onus entirely on a clinician or provider.” As a result, the new questions are similar to the old model, but they are more objective and extensive. The work was supported by a gift from the Argosy Foundation	Clinician, peer, end-user	Health apps, particularly mental health	This framework is broken down into six categories. App Origin and Functionality Inputs and Outputs Privacy and Security Clinical Foundation Features and Engagement Interoperability and Data Sharing 105 Questions

Appendix B. Framework Criteria

Table B-1. Criteria abstracted from the existing frameworks

Framework Source	Category	Category Description	Question	Response Options
uMARS ¹	Engagement	Fun, interesting, customisable, interactive, has prompts (e.g. Sends alerts, messages, reminders, feedback, enables sharing)	Entertainment: Is the app fun/entertaining to use? Does it have components that make it more fun than other similar apps?	1 Dull, not fun or entertaining at all 2 Mostly boring 3 OK, fun enough to entertain user for a brief time (< 5 minutes) 4 Moderately fun and entertaining, would entertain user for some time (5-10 minutes total) 5 Highly entertaining and fun, would stimulate repeat use
uMARS ¹	Engagement	Fun, interesting, customisable, interactive, has prompts (e.g. Sends alerts, messages, reminders, feedback, enables sharing)	Interest: Is the app interesting to use? Does it present its information in an interesting way compared to other similar apps?	1 Not interesting at all 2 Mostly uninteresting 3 OK, neither interesting nor uninteresting; would engage user for a brief time (< 5 minutes) 4 Moderately interesting; would engage user for some time (5-10 minutes total) 5 Very interesting, would engage user in repeat use
uMARS ¹	Engagement	Fun, interesting, customisable, interactive, has prompts (e.g. Sends alerts, messages, reminders, feedback, enables sharing)	Customisation: Does it allow you to customise the settings and preferences that you would like to (e.g. sound, content and notifications)?	1 Does not allow any customisation or requires setting to be input every time 2 Allows little customisation and that limits app's functions 3 Basic customisation to function adequately 4 Allows numerous options for customisation 5 Allows complete tailoring the user's characteristics/preferences, remembers all settings
uMARS ¹	Engagement	Fun, interesting, customisable, interactive, has prompts (e.g. Sends alerts, messages, reminders, feedback, enables sharing)	Interactivity: Does it allow user input, provide feedback, contain prompts (reminders, sharing options, notifications, etc.)?	1 No interactive features and/or no response to user input 2 Some, but not enough interactive features which limits app's functions 3 Basic interactive features to function adequately 4 Offers a variety of

Framework Source	Category	Category Description	Question	Response Options
				interactive features, feedback and user input options 5 Very high level of responsiveness through interactive features, feedback and user input options
uMARS ¹	Engagement	Fun, interesting, customisable, interactive, has prompts (e.g. Sends alerts, messages, reminders, feedback, enables sharing)	Target group: Is the app content (visuals, language, design) appropriate for the target audience	1 Completely inappropriate, unclear or confusing 2 Mostly inappropriate, unclear or confusing 3 Acceptable but not specifically designed for the target audience. May be inappropriate/ unclear/confusing at times 4 Designed for the target audience, with minor issues 5 Designed specifically for the target audience, no issues found
uMARS ¹	Functionality	App functioning, easy to learn, navigation, flow logic, and gestural design of app	Performance: How accurately/fast do the app features (functions) and components (buttons/menus) work?	1 App is broken; no/insufficient/inaccurate response (e.g. crashes/bugs/broken features, etc.) 2 Some functions work, but lagging or contains major technical problems 3 App works overall. Some technical problems need fixing, or is slow at times 4 Mostly functional with minor/negligible problems 5 Perfect/timely response; no technical bugs found, or contains a 'loading time left' indicator (if relevant)
uMARS ¹	Functionality	App functioning, easy to learn, navigation, flow logic, and gestural design of app	Ease of use: How easy is it to learn how to use the app; how clear are the menu labels, icons and instructions?	1 No/limited instructions; menu labels, icons are confusing; complicated 2 Takes a lot of time or effort 3 Takes some time or effort 4 Easy to learn (or has clear instructions) 5 Able to use app immediately; intuitive;

Framework Source	Category	Category Description	Question	Response Options
				simple (no instructions needed)
uMARS ¹	Functionality	App functioning, easy to learn, navigation, flow logic, and gestural design of app	Navigation: Does moving between screens make sense; Does app have all necessary links between screens?	<ul style="list-style-type: none"> 1 No logical connection between screens at all /navigation is difficult 2 Understandable after a lot of time/effort 3 Understandable after some time/effort 4 Easy to understand/navigate 5 Perfectly logical, easy, clear and intuitive screen flow throughout, and/or has shortcuts
uMARS ¹	Functionality	App functioning, easy to learn, navigation, flow logic, and gestural design of app	Gestural design: Do taps/swipes/pinches/sc rolls make sense? Are they consistent across all components/screens?	<ul style="list-style-type: none"> 1 Completely inconsistent/confusing 2 Often inconsistent/confusing 3 OK with some inconsistencies/confusing elements 4 Mostly consistent/intuitive with negligible problems 5 Perfectly consistent and intuitive
uMARS ¹	Aesthetics	Graphic design, overall visual appeal, colour scheme, and stylistic consistency	Layout: Is arrangement and size of buttons, icons, menus and content on the screen appropriate?	<ul style="list-style-type: none"> 1 Very bad design, cluttered, some options impossible to select, locate, see or read 2 Bad design, random, unclear, some options difficult to select/locate/see/read 3 Satisfactory, few problems with selecting/locating/seeing/reading items 4 Mostly clear, able to select/locate/see/read items 5 Professional, simple, clear, orderly, logically organised
uMARS ¹	Aesthetics	Graphic design, overall visual appeal, colour scheme, and stylistic consistency	Graphics: How high is the quality/resolution of graphics used for buttons, icons, menus and content?	<ul style="list-style-type: none"> 1 Graphics appear amateur, very poor visual design - disproportionate, stylistically inconsistent 2 Low quality/low resolution graphics; low quality visual design – disproportionate 3 Moderate quality graphics and visual design (generally

Framework Source	Category	Category Description	Question	Response Options
				<p>consistent in style)</p> <p>4 High quality/resolution graphics and visual design – mostly proportionate, consistent in style</p> <p>5 Very high quality/resolution graphics and visual design - proportionate, consistent in style throughout</p>
uMARS ¹	Aesthetics	Graphic design, overall visual appeal, colour scheme, and stylistic consistency	Visual appeal: How good does the app look?	<p>1 Ugly, unpleasant to look at, poorly designed, clashing, mismatched colours</p> <p>2 Bad – poorly designed, bad use of colour, visually boring</p> <p>3 OK – average, neither pleasant, nor unpleasant</p> <p>4 Pleasant – seamless graphics – consistent and professionally designed</p> <p>5 Beautiful – very attractive, memorable, stands out; use of colour enhances app features/menus</p>
uMARS ¹	Information	Contains high quality information (e.g. Text, feedback, measures, references) from a credible source	Quality of information: Is app content correct, well written, and relevant to the goal/topic of the app?	<p>N/A There is no information within the app</p> <p>1 Irrelevant/ inappropriate/ incoherent/incorrect</p> <p>2 Poor. Barely relevant/appropriate/coherent/may be incorrect</p> <p>3 Moderately relevant/appropriate/coherent/and appears correct</p> <p>4 Relevant/ appropriate/ coherent/correct</p> <p>5 Highly relevant, appropriate, coherent, and correct</p>
uMARS ¹	Information	Contains high quality information (e.g. Text, feedback, measures, references) from a credible source	Quantity of information: Is the information within the app comprehensive but concise?	<p>N/A There is no information within the app</p> <p>1 Minimal or overwhelming</p> <p>2 Insufficient or possibly overwhelming</p> <p>3 OK but not comprehensive or concise</p>

Framework Source	Category	Category Description	Question	Response Options
				4 Offers a broad range of information, has some gaps or unnecessary detail; or has no links to more information and resources 5 Comprehensive and concise; contains links to more information and resources
uMARS ¹	Information	Contains high quality information (e.g. Text, feedback, measures, references) from a credible source	Visual information: Is visual explanation of concepts – through charts/graphs/images/v ideas, etc. – clear, logical, correct?	N/A There is no visual information within the app (e.g. it only contains audio, or text) 1 Completely unclear/confusing/wrong or necessary but missing 2 Mostly unclear/confusing/wrong 3 OK but often unclear/confusing/wrong 4 Mostly clear/logical/correct with negligible issues 5 Perfectly clear/logical/correct
uMARS ¹	Information	Contains high quality information (e.g. Text, feedback, measures, references) from a credible source	Credibility of source: does the information within the app seem to come from a credible source?	N/A There is no information within the app 1 Suspicious source 2 Lacks credibility 3 Not suspicious but legitimacy of source is unclear 4 Possibly comes from a legitimate source 5 Definitely comes from a legitimate/specialised source
uMARS ¹	Quality	App subjective quality	Would you recommend this app to people who might benefit from it?	1 Not at all I would not recommend this app to anyone 2 There are very few people I would recommend this app to 3 Maybe There are several people I would recommend this app to 4 There are many people I would recommend this app to 5 Definitely I would recommend this app to everyone
uMARS ¹	Quality	App subjective quality	How many times do you think you would	1 None 2 1-2

Framework Source	Category	Category Description	Question	Response Options
			use this app in the next 12 months if it was relevant to you?	3 3-10 4 10-50 5 >50
uMARS ¹	Quality	App subjective quality	Would you pay for this app?	1 Definitely not 2 3 4 5 Definitely yes
uMARS ¹	Quality	App subjective quality	What is your overall (star) rating of the app?	1 * One of the worst apps I've used 2 ** 3 *** Average 4 **** 5 ***** One of the best apps I've used
uMARS ¹	Perceived impact	NA	Awareness – This app has increased my awareness of the importance of addressing the health behaviour	1. Strongly disagree 2 3 4 5. Strongly Agree
uMARS ¹	Perceived impact	NA	Knowledge – This app has increased my knowledge/understanding of the health behaviour	1. Strongly disagree 2 3 4 5. Strongly Agree
uMARS ¹	Perceived impact	NA	Attitudes – The app has changed my attitudes toward improving this health behaviour	1. Strongly disagree 2 3 4 5. Strongly Agree
uMARS ¹	Perceived impact	NA	Intention to change – The app has increased my intentions/motivation to address this health behavior	1. Strongly disagree 2 3 4 5. Strongly Agree
uMARS ¹	Perceived impact	NA	Help seeking – This app would encourage me to seek further help to address the health behavior (if I needed it)	1. Strongly disagree 2 3 4 5. Strongly Agree
uMARS ¹	Perceived impact	NA	Behaviour change – Use of this app will increase/decrease the health behaviour	1. Strongly disagree 2 3 4 5. Strongly Agree"
MIND ²	App Origin	NA	Does it come from the government?	NA
MIND ²	App Origin	NA	Does it come from a for-profit company or developer?	NA
MIND ²	App Origin	NA	Does it come from a non-profit company?	NA
MIND ²	App Origin	NA	Does it come from a	NA

Framework Source	Category	Category Description	Question	Response Options
			trusted healthcare company?	
MIND ²	App Origin	NA	Does it come from an academic institution?	NA
MIND ²	App Functionality	NA	Does it work on Apple(iOS)?	NA
MIND ²	App Functionality	NA	What is the Apple version?	NA
MIND ²	App Functionality	NA	What is the oldest iOS version supported?	NA
MIND ²	App Functionality	NA	What was the Apple release date?	NA
MIND ²	App Functionality	NA	When was the last Apple (IOS) update?	NA
MIND ²	App Functionality	NA	Has the apple version been updated in the last 180 days?	NA
MIND ²	App Functionality	NA	Number of reviews on Apple store?	NA
MIND ²	App Functionality	NA	Rating (number of stars) on Apple store?	NA
MIND ²	App Functionality	NA	App size on iOS?	NA
MIND ²	App Functionality	NA	Does it work on Android?	NA
MIND ²	App Functionality	NA	What is the Android version?	NA
MIND ²	App Functionality	NA	What is the oldest Android version supported?	NA
MIND ²	App Functionality	NA	What was the Google play store release date?	NA
MIND ²	App Functionality	NA	When was the last Android update?	NA
MIND ²	App Functionality	NA	Has the android version been updated in the last 180 days?	NA
MIND ²	App Functionality	NA	Number of reviews on google play store?	NA
MIND ²	App Functionality	NA	Rating (number of stars) on google play store?	NA
MIND ²	App Functionality	NA	App size on android?	NA
MIND ²	App Functionality	NA	Does the app work offline?	NA
MIND ²	App Functionality	NA	Does it have at least one accessibility feature (like adjust text size, text to voice, or	NA
MIND ²	App Functionality	NA	colorblind color scheme adjuster)?	NA
MIND ²	App	NA	Does it have a web	NA

Framework Source	Category	Category Description	Question	Response Options
	Functionality		version?	
MIND ²	App Functionality	NA	Does it work with Spanish?	NA
MIND ²	App Functionality	NA	Does it work with a language other than English or Spanish?	NA
MIND ²	App Functionality	NA	Is the app totally free?	NA
MIND ²	App Functionality	NA	What is the cost up front?	NA
MIND ²	App Functionality	NA	Are there in-app purchases?	NA
MIND ²	App Functionality	NA	Is it a subscription (recurrent/monthly/annual)?	NA
MIND ²	Input	NA	Input: surveys?	NA
MIND ²	Input	NA	Input: Diary?	NA
MIND ²	Input	NA	Input: Geolocation?	NA
MIND ²	Input	NA	Input: contact list?	NA
MIND ²	Input	NA	Input: Camera?	NA
MIND ²	Input	NA	Input: Microphone?	NA
MIND ²	Input	NA	Input: step count?	NA
MIND ²	Input	NA	Input: external devices (e.g. a wearable sending direct data)?	NA
MIND ²	Input	NA	Input: social network?	NA
MIND ²	Output	NA	Output: notifications?	NA
MIND ²	Output	NA	Output: psychoeducational references/information?	NA
MIND ²	Output	NA	Output: social network?	NA
MIND ²	Output	NA	Output: reminders?	NA
MIND ²	Output	NA	Output: graphs of data?	NA
MIND ²	Output	NA	Output: summary of data (in text or numbers)?	NA
MIND ²	Output	NA	Output: link to formal care/coaching?	NA
MIND ²	Privacy & Security	NA	Is there a privacy policy?	NA
MIND ²	Privacy & Security	NA	Does the app declare data use and purpose?	NA
MIND ²	Privacy & Security	NA	Does the app report security measures in place?	NA
MIND ²	Privacy & Security	NA	Is PHI shared?	NA
MIND ²	Privacy & Security	NA	Is de-identified data shared?	NA
MIND ²	Privacy & Security	NA	Is anonymized/aggregate data shared?	NA
MIND ²	Privacy & Security	NA	Can you opt out of data collection?	NA
MIND ²	Privacy &	NA	Can you delete your	NA

Framework Source	Category	Category Description	Question	Response Options
	Security		data?	
MIND ²	Privacy & Security	NA	Is the user data stored only on the device?	NA
MIND ²	Privacy & Security	NA	Is the user data stored on a server?	NA
MIND ²	Privacy & Security	NA	Does the app have a crisis management feature?	NA
MIND ²	Privacy & Security	NA	Does the app claim it meets HIPAA (or analogous national standard for patient/PHI	NA
MIND ²	Privacy & Security	NA	privacy protection)	NA
MIND ²	Privacy & Security	NA	Reading level of the privacy policy (what grade reading level)?	NA
MIND ²	Privacy & Security	NA	Does the app use 3rd party vendors (i.e. google analytics, etc.)?	NA
MIND ²	Evidence & Clinical Foundation	NA	Is the app content well-written, correct, and relevant?	NA
MIND ²	Evidence & Clinical Foundation	NA	Does the app appear to do what it claims to do?	NA
MIND ²	Evidence & Clinical Foundation	NA	Is the app patient facing?	NA
MIND ²	Evidence & Clinical Foundation	NA	How many feasibility/usability studies?	NA
MIND ²	Evidence & Clinical Foundation	NA	What is the highest feasibility impact factor?	NA
MIND ²	Evidence & Clinical Foundation	NA	How many evidence/efficacy studies?	NA
MIND ²	Evidence & Clinical Foundation	NA	What is the highest efficacy impact factor?	NA
MIND ²	Evidence & Clinical Foundation	NA	Can the app cause harm?	NA
MIND ²	Evidence & Clinical Foundation	NA	Does the app provide any warning for use?	NA
MIND ²	Features	NA	Features: mood tracking?	NA
MIND ²	Features	NA	Features: medication tracking?	NA
MIND ²	Features	NA	Features: sleep tracking?	NA
MIND ²	Features	NA	Features: physical	NA

Framework Source	Category	Category Description	Question	Response Options
			exercise tracking?	
MIND ²	Features	NA	Features: psychoeducation?	NA
MIND ²	Features	NA	Features: journaling?	NA
MIND ²	Features	NA	Features: picture gallery/hope board?	NA
MIND ²	Features	NA	Features: mindfulness?	NA
MIND ²	Features	NA	Features: deep breathing?	NA
MIND ²	Features	NA	Features: iCBT or sleep therapy?	NA
MIND ²	Features	NA	Features: CBT?	NA
MIND ²	Features	NA	Features: ACT?	NA
MIND ²	Features	NA	Features: DBT?	NA
MIND ²	Features	NA	Features: peer support?	NA
MIND ²	Features	NA	Features: connection to coach/therapist?	NA
MIND ²	Features	NA	Features: biodata?	NA
MIND ²	Features	NA	Features: goal setting/habits?	NA
MIND ²	Features	NA	Features: physical health exercises?	NA
MIND ²	Features	NA	Features: Chatbot interaction (like with virtual character)?	NA
MIND ²	Features	NA	Features: Biofeedback with sense data (e.g., HRV, skin conductance, etc)?	NA
MIND ²	Engagement Style	NA	Engagement style: user generated data?	NA
MIND ²	Engagement Style	NA	Engagement style: chat/message based?	NA
MIND ²	Engagement Style	NA	Engagement style: is it a screener/assessment?	NA
MIND ²	Engagement Style	NA	Engagement style: real time response?	NA
MIND ²	Engagement Style	NA	Engagement style: Asynchronous response?	NA
MIND ²	Engagement Style	NA	Engagement style: gamification (points, badges)?	NA
MIND ²	Engagement Style	NA	Engagement style: videos?	NA
MIND ²	Engagement Style	NA	Engagement style: audio/music/scripts?	NA
MIND ²	Engagement Style	NA	Engagement style: AI support?	NA
MIND ²	Engagement Style	NA	Engagement style: peer support?	NA
MIND ²	Engagement Style	NA	Engagement style: network support?	NA

Framework Source	Category	Category Description	Question	Response Options
MIND ²	Engagement Style	NA	Engagement style: Collaborative with provider/other?	NA
MIND ²	App Use	NA	Is it a self-help/self-management tool?	NA
MIND ²	App Use	NA	Is it a reference app?	NA
MIND ²	App Use	NA	Is it intended for hybrid use with a clinician in conjunction with treatment plan?	NA
MIND ²	Interoperability & Data Sharing	NA	Do you own your data?	NA
MIND ²	Interoperability & Data Sharing	NA	Can you email or export your data?	NA
MIND ²	Interoperability & Data Sharing	NA	Can you send your data to a medical record?	NA
Health Navigator-NZ ³	User experience – is the app user friendly?	NA	Is the layout simple, clear and well designed? Or cluttered and confusing?	NA
Health Navigator-NZ ³	User experience – is the app user friendly?	NA	Is it easy and intuitive to learn how to use? Does it have long lags or technical bugs?	NA
Health Navigator-NZ ³	User experience – is the app user friendly?	NA	Is the app's language and information suitable for the target group your patient belongs to? If local, does it include te reo Māori and other language options?	NA
Health Navigator-NZ ³	User experience – is the app user friendly?	NA	How much data space will the app take up on a smartphone or tablet and/or how much mobile data will it require to run?	NA
Health Navigator-NZ ³	User experience – is the app user friendly?	NA	Does it require the internet to use its core features?	NA
Health Navigator-NZ ³	User experience – is the app user friendly?	NA	Does it have ongoing costs or charges?	NA
Health Navigator-NZ ³	User experience – is the app user friendly?	NA	Does it have advertising? Is the advertising intrusive and distracting?	NA
Health Navigator-NZ ³	User experience – is the app user friendly?	NA	What are the reviews and ratings of the app on sites that you trust? Generally in app	NA

Framework Source	Category	Category Description	Question	Response Options
			stores, review scores are aggregated and used to determine an overall score for the app. Usually the higher the score an app gets, the more people liked it. This is often fuel for the thriving trade in fake reviews.	
Health Navigator-NZ ³	Credibility – is the app credible and safe?	NA	Does the app provide accurate information?	NA
Health Navigator-NZ ³	Credibility – is the app credible and safe?	NA	Does the app make suggestions about changing medication or treatment plans without consultation with the person's health professional?	NA
Health Navigator-NZ ³	Credibility – is the app credible and safe?	NA	Is the medical content or advice offered by the app sound, safe and up-to-date?	NA
Health Navigator-NZ ³	Credibility – is the app credible and safe?	NA	Does the app have the ability to handle 'dangerous' information entered by a patient, such as a low blood glucose level or suicidal thoughts?	NA
Health Navigator-NZ ³	Assess the source (the most reliable health apps tend to result from collaborations among developers and health professionals)	NA	Does the app come from a legitimate source?	NA
Health Navigator-NZ ³	Assess the source (the most reliable health apps tend to result from collaborations among developers and health professionals)	NA	Has it had input from health professionals?	NA
Health Navigator-	Assess the source (the	NA	Is it sponsored or developed by a	NA

Framework Source	Category	Category Description	Question	Response Options
NZ ³	most reliable health apps tend to result from collaborations among developers and health professionals)		reputable organisation, university or health provider?	
Health Navigator-NZ ³	Assess the source (the most reliable health apps tend to result from collaborations among developers and health professionals)	NA	Consider its New Zealand relevance	NA
Health Navigator-NZ ³	Assess the source (the most reliable health apps tend to result from collaborations among developers and health professionals)	NA	Is the app suitable or relevant for New Zealanders?	NA
Health Navigator-NZ ³	Assess the source (the most reliable health apps tend to result from collaborations among developers and health professionals)	NA	Most apps available from app stores are developed overseas and don't have a New Zealand focus. It's important that the recommendations within these apps are in keeping with New Zealand practice. Other aspects to look for are things like the option for metric measurements (e.g., kilograms instead of pounds) and medications that may not be available in New Zealand.	NA
Health Navigator-NZ ³	Assess the source (the most reliable health apps tend to result from	NA	Consider privacy and security concerns	NA

Framework Source	Category	Category Description	Question	Response Options
	collaborations among developers and health professionals)			
Health Navigator-NZ ³	Assess the source (the most reliable health apps tend to result from collaborations among developers and health professionals)	NA	Does the app have clear privacy guidelines on how data shared via the app will be stored and used?	NA
Health Navigator-NZ ³	Assess the source (the most reliable health apps tend to result from collaborations among developers and health professionals)	NA	Does the app ask for permission to access unrelated information that may be used for advertising or other commercial purposes?	NA
Health Navigator-NZ ³	Assess the source (the most reliable health apps tend to result from collaborations among developers and health professionals)	NA	Does the app require your credit card details before you begin using it?	NA
Health Navigator-NZ ³	General	NA	Provide health information and advice.	NA
Health Navigator-NZ ³	General	NA	Track your personal health information, such as diet and physical activity.	NA
Health Navigator-NZ ³	General	NA	Perform calculations and analyses, such as calculating your risk of stroke or analysing an image of a skin lesion.	NA
Health Navigator-NZ ³	General	NA	Link to social media, e.g. sharing and comparing running times.	NA
Health Navigator-	General	NA	Provide entertainment for health purposes,	NA

Framework Source	Category	Category Description	Question	Response Options
NZ ³			e.g. carbohydrate counting games for people with diabetes.	
Health Navigator-NZ ³	Privacy and Security	NA	How your personal information could be at risk?	NA
Health Navigator-NZ ³	Privacy and Security	NA	How to improve your safety and security when using apps	NA
Health Navigator-NZ ³	Privacy and Security	NA	What is a privacy policy?	NA
Health Navigator-NZ ³	Privacy and Security	NA	Is your data being shared with a third party or advertiser?	NA
Health Navigator-NZ ³	Privacy and Security	NA	Where is your data stored – on your device or in the cloud?	NA
APA ⁴	Access and Background	NA	Does the app identify ownership?	NA
APA ⁴	Access and Background	NA	Does the app identify funding sources and conflicts of interest?	NA
APA ⁴	Access and Background	NA	Does the app come from a trusted source?	NA
APA ⁴	Access and Background	NA	Does it claim to be medical?	NA
APA ⁴	Access and Background	NA	Are there additional or hidden costs?	NA
APA ⁴	Access and Background	NA	Does the app work offline?	NA
APA ⁴	Access and Background	NA	On which platforms/operating systems does it work?	NA
APA ⁴	Access and Background	NA	Does it work on a desktop computer?	NA
APA ⁴	Access and Background	NA	Does the app work with accessibility features of the iPhone/android?	NA
APA ⁴	Access and Background	NA	Is it accessible for those with impaired vision or other disabilities?	NA
APA ⁴	Access and Background	NA	Has the app been updated in the last 180 days?	NA
APA ⁴	Privacy and Security	NA	Is there a transparent privacy policy that is clear and accessible before use?	NA
APA ⁴	Privacy and Security	NA	Does the app declare data use and purpose?	NA
APA ⁴	Privacy and Security	NA	Does the app describe use of PHI?	NA
APA ⁴	Privacy and	NA	Deidentified vs.	NA

Framework Source	Category	Category Description	Question	Response Options
	Security		anonymous?	
APA ⁴	Privacy and Security	NA	Can you opt out of data collection or delete data?	NA
APA ⁴	Privacy and Security	NA	Are data maintained in the device or on the web?	NA
APA ⁴	Privacy and Security	NA	Does the app explain security systems used?	NA
APA ⁴	Privacy and Security	NA	Does the app collect, use, and/or transmit sensitive data? If yes, does it claim to do so securely?	NA
APA ⁴	Privacy and Security	NA	What third parties does the app share data with?	NA
APA ⁴	Privacy and Security	NA	If appropriate, is the app equipped to respond to potential harms or safety concerns?	NA
APA ⁴	Clinical Foundation	NA	Does the app appear to do what it claims to do?	NA
APA ⁴	Clinical Foundation	NA	Is the app content correct, well-written, and relevant?	NA
APA ⁴	Clinical Foundation	NA	What are the relevant sources or references supporting the app use cases?	NA
APA ⁴	Clinical Foundation	NA	Is there evidence of specific benefit from academic institutions, publications, end user feedback, or research studies?	NA
APA ⁴	Clinical Foundation	NA	Is there evidence of effectiveness/efficacy?	NA
APA ⁴	Clinical Foundation	NA	Was there an attempt to validate app usability and feasibility?	NA
APA ⁴	Clinical Foundation	NA	Does the app have a clinical/recovery foundation relevant to your intended use?	NA
APA ⁴	Usability	NA	What are the main engagement styles of the app?	NA
APA ⁴	Usability	NA	Do the app and its features align with your needs and priorities?	NA
APA ⁴	Usability	NA	Is it customizable?	NA
APA ⁴	Usability	NA	Does the app clearly define functional	NA

Framework Source	Category	Category Description	Question	Response Options
			scope?	
APA ⁴	Usability	NA	Does the app seem easy to use?	NA
APA ⁴	Data Integration towards Therapeutic Goal	NA	Do you own your data?	NA
APA ⁴	Data Integration towards Therapeutic Goal	NA	Can data be easily shared and interpreted in a way that's consistent with the stated purpose of the app?	NA
APA ⁴	Data Integration towards Therapeutic Goal	NA	Can the app share data with EMR and other data tools (apple Healthkit, FitBit)?	NA
APA ⁴	Data Integration towards Therapeutic Goal	NA	Is the app for individual use or to be used in collaboration with a provider?	NA
APA ⁴	Data Integration towards Therapeutic Goal	NA	If intended to be used with a provider, does the app have the ability to export or transfer data?	NA
APA ⁴	Data Integration towards Therapeutic Goal	NA	Does the app lead to any positive behavior change or skill acquisition?	NA
APA ⁴	Data Integration towards Therapeutic Goal	NA	Does the app improve therapeutic alliance between patient and provider?	NA
One Mind Psyber Guide ⁵	Credibility	The Credibility Score combines information about research, development, purpose, and popularity. This measure aims to give users an idea of how credible a digital tool is, i.e. How likely it is that it will work. Apps are scored based on:	Consumer Ratings	2 Ratings exist from >1500 users with an average rating of 3.5+ 1 Ratings exist from 31-1500 users with an average rating of 3.5+ 0 Fewer than 30 user rating OR an average rating below 3.5
One Mind Psyber Guide ⁵	Credibility	The Credibility Score combines information about research, development, purpose, and popularity. This measure aims to give	Proposed Goal	2 Product describes at least one mental health goal which is specific, measurable, and achievable (e.g. reduce stress, reduce symptoms

Framework Source	Category	Category Description	Question	Response Options
		users an idea of how credible a digital tool is, i.e. How likely it is that it will work. Apps are scored based on:		of PTSD) 1 Product describes non-specific or hard to measure mental health goals (e.g. improve your life, improve your wellbeing) 0 No clear goals
One Mind Psyber Guide ⁵	Credibility	The Credibility Score combines information about research, development, purpose, and popularity. This measure aims to give users an idea of how credible a digital tool is, i.e. How likely it is that it will work. Apps are scored based on:	Evidence-Based Content	1 The app uses evidence-based practices to achieve its goals 0 The app does not use evidence-based practices to achieve its goals (or there are no goals described)
One Mind Psyber Guide ⁵	Credibility	The Credibility Score combines information about research, development, purpose, and popularity. This measure aims to give users an idea of how credible a digital tool is, i.e. How likely it is that it will work. Apps are scored based on:	Research Base	3 Strong research support for the product (at least two between-group design experiments that show efficacy or effectiveness) 2 Some research support for the product (at least one experiment that shows efficacy or effectiveness) 1 Other research (e.g. single case designs, quasi-experimental methods demonstrating efficacy, or preliminary analyses) 0 No research.
One Mind Psyber Guide ⁵	Credibility	The Credibility Score combines information about research, development, purpose, and popularity. This measure aims to give users an idea of how credible a digital tool is, i.e. How likely it is that it will work. Apps are scored based on:	Software Updates	2 The application has been revised within the last 6 months 1 The application has been revised within the last 12 months 0 The application has not been revised or was revised more than 12 months ago.
One Mind Psyber Guide ⁵	Credibility	The Credibility Score combines information about research, development, purpose, and popularity. This measure aims to give users an idea of how credible a digital tool is,	Clinical Input in Development	1 Clinical leader with mental health expertise involved in development 0 No clinical leader with mental health expertise involved in development

Framework Source	Category	Category Description	Question	Response Options
		i.e. How likely it is that it will work. Apps are scored based on:		
One Mind Psyber Guide ⁵	Credibility	The Credibility Score combines information about research, development, purpose, and popularity. This measure aims to give users an idea of how credible a digital tool is, i.e. How likely it is that it will work. Apps are scored based on:	Research on Development Process	1 Pilot, feasibility and acceptability data OR evidence of stakeholder engagement in development 0 No pilot, feasibility and acceptability data AND no evidence of stakeholder engagement
One Mind Psyber Guide ⁵	Credibility	The Credibility Score combines information about research, development, purpose, and popularity. This measure aims to give users an idea of how credible a digital tool is, i.e. How likely it is that it will work. Apps are scored based on:	Efficacy of Other Products	1 Developer/development team has developed other mental health interventions delivered via technological medium which demonstrate efficacy 0 No other mental health technological interventions demonstrating efficacy have been developed by this team
One Mind Psyber Guide ⁵	Credibility	The Credibility Score combines information about research, development, purpose, and popularity. This measure aims to give users an idea of how credible a digital tool is, i.e. How likely it is that it will work. Apps are scored based on:	Research Independence & Review	0 At least one research paper funded by government agency (e.g. NIH) or non-profit organization OR two articles published in peer-reviewed journals 1 All research funded primarily by for-profit organizations or combined funding sources OR one article published in a peer-reviewed journal 2 No information about source of funding for the research AND No published, peer-reviewed papers
One Mind Psyber Guide ⁵	User Experience	"User Experience", sometimes referred to as just UX, is the overall experience of using an app or program, in terms of how easy and engaging it is to use. The Mobile App Rating Scale (MARS) is used to assess	The MARS mean is the mean of four objective subscales: Engagement: how fun, interesting and customizable the app is, and how well it engages the people it's intended for	The Subjective Quality and Perceived Impact scores are based on the raters' own impression of the eTool, including its usability and perceived effectiveness. The MARS can be used as an adjunct to

Framework Source	Category	Category Description	Question	Response Options
		<p>the quality of the user experience of apps. MARS was developed by a team of researchers at Queensland University of Technology (QUT), with expertise in the development of digital health tools</p>	<p>Functionality: how well the app features work, how easy it is to navigate through the app. Is it self-explanatory, intuitive, and easy to learn? Aesthetics: the overall visual design – how appealing are the graphics, colors and layout? Information: is the content of the app accurate, well-written and credible? Subjective Quality Perceived Impact</p>	<p>qualitative eTool descriptions, to give eTool users an overview of their quality rating. The scale can also help with the ranking of eTools based on their quality. The MARS scale is being used worldwide by eTool evaluation and development projects.</p>
One Mind Psyber Guide ⁵	Transparency	<p>Transparency scores relate to information regarding an apps' data storage and collection policies and how readily available this information is to users. It's important to note here that for this metric, we evaluate whether or not an app's privacy policy has certain key pieces of information regarding data storage, encryption, deletion, etc. What we don't do is audit the apps practices, to ensure that they actually do what they say they do in their policies. We believe that developers should be as transparent as possible with privacy information so that users can be fully informed of how their data is used and stored.</p>	Transparency	<p>Acceptable A product that has been scored as acceptable has an acceptable level of data transparency; the privacy policy of the product provides sufficient and easily accessible information on the policies related to data collection, storage, and exchange. The information provided conforms to standards for collection, storage, and exchange of health information. Questionable A product that has been scored as questionable has a privacy policy that is unclear or lacking specific details of policies surrounding data collection, storage, and exchange or is questionable in its adherence to standards on collection, storage, and exchange of health information. Unacceptable A product that has been scored as unacceptable either a) does not have a privacy policy, b) has a privacy policy that excludes important information</p>

Framework Source	Category	Category Description	Question	Response Options
				about data privacy, collection, storage, or exchange, or c) has a privacy policy that outlines practices for data privacy, collection, storage or exchange that do not conform to standards for health information.
Martinengo et al., 2019 ⁶	Features	NA	Tracking of mood	NA
Martinengo et al., 2019 ⁶	Features	NA	Development of a safety plan	NA
Martinengo et al., 2019 ⁶	Features	NA	Recommendation of activities to prevent suicidal thoughts	NA
Martinengo et al., 2019 ⁶	Features	NA	Information and Education	NA
Martinengo et al., 2019 ⁶	Features	NA	Access to support networks	NA
Martinengo et al., 2019 ⁶	Features	NA	Access to emergency counseling	NA
Martinengo et al., 2019 ⁶	Features	NA	Trustworthiness of the information provided by app	NA
Moshi, et al., 2018 ⁷	Evidence	NA	Diagnostic Accuracy	NA
Moshi, et al., 2018 ⁷	Evidence	NA	Therapeutic Effectiveness	NA
Moshi, et al., 2018 ⁷	Organizational	NA	Organizational Aspects - - whether the clinician would use it daily, what kind of training would be required.	NA
Moshi, et al., 2018 ⁷	Legal & Ethical	NA	Legal	NA
Moshi, et al., 2018 ⁷	Legal & Ethical	NA	Ethical	NA
Webb, et al., 2010 ⁸	Features	NA	Behavior Change Techniques	NA
Ondersma, et al., 2020 ⁹	Evidence	NA	Statistical Analysis	NA
UK Digital Health Framework ¹⁰	Company information	NA	Provide the name of your company	Free text
UK Digital Health Framework ¹⁰	Company information	NA	Provide the name of your product	Free text
UK Digital Health Framework ¹⁰	Company information	NA	Provide the type of product	App Wearable Software as a Service (SaaS) Other
UK Digital Health Framework ¹⁰	Company information	NA	Provide the name and job title of the individual who will be the key	Free text

Framework Source	Category	Category Description	Question	Response Options
			contact at your organisation	
UK Digital Health Framework ¹⁰	Company information	NA	Provide the key contact's email address	Free text
UK Digital Health Framework ¹⁰	Company information	NA	Provide the key contact's phone number	Free text
UK Digital Health Framework ¹⁰	Company information	NA	Provide the registered address of your company	Free text
UK Digital Health Framework ¹⁰	Company information	NA	In which country is your organization registered?	Free text
UK Digital Health Framework ¹⁰	Company information	NA	If you have a Companies House registration in the UK please provide your number	Free text
UK Digital Health Framework ¹⁰	Company information	NA	If applicable, when was your last assessment from the Care Quality Commission (CQC)?	Date Not applicable
UK Digital Health Framework ¹⁰	Company information	NA	If applicable, upload your latest CQC report.	Upload
UK Digital Health Framework ¹⁰	Value proposition	NA	Who is this product intended to be used for?	Patients Diagnostics Clinical Support Infrastructure Workforce Other
UK Digital Health Framework ¹⁰	Value proposition	This question is a context question and therefore a high level summary is required.	Provide a clear description of what the product is designed to do and of how it is expected to be used	Free text
UK Digital Health Framework ¹⁰	Value proposition	This question is a context question and therefore a high level summary is required. If your product has had an evaluation or undergone clinical trials include this information.	Describe clearly the intended or proven benefits for users and confirm if / how the benefits have been validated	Free text
UK Digital Health Framework ¹⁰	Value proposition	This question is a context question and it is expected that existing documentation will be provided.	Please attach one or more user journeys which were used in the development of this product	Attached Not available
UK Digital Health Framework ¹⁰	Value proposition	GOV.UK provides guidance on how to make a user journey map and what should be included. Data flows enable the assessor to understand how data moves through a product. This may be	Where possible please also provide your data flows	NA

Framework Source	Category	Category Description	Question	Response Options
		included within a Data Protection Impact Assessment. If this is the case, please provide as a separate attachment for ease of review.		
UK Digital Health Framework ¹⁰	Clinical Safety	<p>The DCB0129 standard applies to organizations that are responsible for the development and maintenance of health IT systems. A health IT system is defined as “product used to provide electronic information for health and social care purposes”.</p> <p>To pass, the developer is required to confirm that they have undertaken Clinical Risk Management activities in compliance with DCB0129.</p>	Have you undertaken Clinical Risk Management activities for this product which comply with DCB0129?	Yes No
UK Digital Health Framework ¹⁰	Clinical Safety	<p>DCB0129 sets out the activities that must and should be undertaken for health IT systems. An example clinical risk management system template can be downloaded from the NHS Digital website. To pass, the developer is required to evidence that a clinical risk management system is in place and that it is compliant with the requirements set out in DCB0129.</p> <p>This should include:</p> <ul style="list-style-type: none"> ● The clinical risk management governance arrangements that are in place ● The clinical risk management activities ● Clinical safety competence and training ● Audits <p>Attached No evidence available Specifically your DTAC submission should include:</p>	Please supply your Clinical Safety Case Report and Hazard Log	Attached No evidence attached

Framework Source	Category	Category Description	Question	Response Options
		<ul style="list-style-type: none"> ● A summary of the product and its intended use ● A summary of clinical risk management activities ● A summary of hazards identified which you have been unable to mitigate to as low as it is reasonably practicable ● The clear identification of hazards which will require user or commissioner action to reach acceptable mitigation (for example, training and business process change) <p>It should not include the hazard log in the body of the document - this should be supplied separately.</p> <p>Example Clinical Safety Case Report and Hazard Log templates can be downloaded from the NHS Digital website.</p> <p>To pass, the developer is required to submit the Clinical Safety Case Report and Hazard Log that is compliant with the requirements set out in DCB0129. This should be commensurate with the scale and clinical functionality of the product and address the clinical risk management activities specified with the standard.</p> <p>The Clinical Safety Case Report should present the arguments and supporting evidence that provides a compelling, comprehensible and valid case that a system is safe for a given application in a given environment at the defined point in the</p>		

Framework Source	Category	Category Description	Question	Response Options
		<p>products lifecycle.</p> <p>It should provide the reader with a summary of all the relevant knowledge that has been acquired relating to the clinical risks associated with the product at that point in the life cycle:</p> <ul style="list-style-type: none"> ● A clear and concise record of the process that has been applied to determine the clinical safety of the product ● A summary of the outcomes of the assessment procedures applied ● A clear listing of any residual clinical risks that have been identified and the related operational constraints and limitations that are applicable ● A clear listing of any hazards and associated clinical risks that have been transferred, together with any declared risk control measures, that are to be addressed as part of the clinical risk management process in the organisation where the product is being deployed ● A listing of outstanding test issues / defects associated with the product which may have a clinical safety impact. <p>The Hazard Log should record and communicate the on-going identification and resolution of hazards associated with the product. All foreseeable hazards should be identified and the risk of such hazards should be reduced to acceptable levels.</p>		

Framework Source	Category	Category Description	Question	Response Options
		<p>A summary should also be provided to the assessor of identified hazards that the developer has been unable to mitigate to as low as it is reasonably practicable. It should also clearly identify the hazards which will require user or commissioner action to reach acceptable mitigation.</p>		
UK Digital Health Framework ¹⁰	Clinical Safety	<p>The CSO must:</p> <ul style="list-style-type: none"> • Be a suitably qualified and experienced clinician • Hold a current registration with an appropriate professional body relevant to their training and experience • Be knowledgeable in risk management and its application to clinical domains • Be suitably trained and qualified in risk management or have an understanding in principles of risk and safety as applied to Health IT • Have completed appropriate training <p>The work of the CSO can be undertaken by an outsourced third party. To pass, the developer must have a named CSO which can be through an outsourced arrangement. They must be a suitably qualified and experienced clinician and hold a current registration with an appropriate professional body relevant to their training and experience.</p>	Please provide the name of your Clinical Safety Officer (CSO), their profession and registration details	Free text
UK Digital Health Framework ¹⁰	Clinical Safety	If this question is not applicable, because your product does not fall within the UK Medical Devices Regulations 2002, continue to question C1.4.	If your product falls within the UK Medical Devices Regulations 2002, is it registered with the Medicines and Healthcare products Regulatory Agency	Yes No Not applicable

Framework Source	Category	Category Description	Question	Response Options
			(MHRA)?	
UK Digital Health Framework ¹⁰	Clinical Safety	NA	If the UK Medical Device Regulations 2002 are applicable, please provide your Declaration of Conformity and, if applicable, certificate of conformity issued by a Notified Body / UK Approved Body	NA
UK Digital Health Framework ¹⁰	Clinical Safety	NA	Do you use or connect to any third party products? If yes, please attach relevant Clinical Risk Management documentation and conformity certificate	Yes No
UK Digital Health Framework ¹⁰	Data Protection	NA	If you are required to register with the Information Commissioner please attach evidence of a current registration.	Attached Not provided
UK Digital Health Framework ¹⁰	Data Protection	NA	If you are not required to register please attach a completed self-assessment showing the outcome from the Information Commissioner and your responses which support this determination.	Yes No We do not need one
UK Digital Health Framework ¹⁰	Data Protection	NA	Do you have a nominated Data Protection Officer (DPO)? If you are required to have a nominated Data Protection Officer, please provide their name. If you are not required to have a DPO please attach a completed self assessment showing the outcome from the Information Commissioner and your responses which support this determination. If you are not required to have a DPO please	Free text Attachment

Framework Source	Category	Category Description	Question	Response Options
			attach a completed self assessment showing the outcome from the Information Commissioner and your responses which support this determination.	
UK Digital Health Framework ¹⁰	Data Protection	NA	Does your product have access to any personally identifiable data or NHS held patient data?	Yes No
UK Digital Health Framework ¹⁰	Data Protection	NA	Please confirm you are compliant (having standards met or exceeded status) with the annual Data Security and Protection Toolkit Assessment.	Confirmed Unable to confirm
UK Digital Health Framework ¹⁰	Data Protection	NA	If you have not completed the current year's assessment and the deadline has not yet passed, please confirm that you intend to complete this ahead of the deadline and that there are no material changes from your previous years submission that would affect your compliance.	Attached Not provided
UK Digital Health Framework ¹⁰	Data Protection	NA	Please attach the Data Protection Impact Assessment (DPIA) relating to the product.	Confirm Cannot confirm
UK Digital Health Framework ¹⁰	Data Protection	NA	Please confirm your risk assessments and mitigations / access controls / system level security policies have been signed-off by your Data Protection Officer (if one is in place) or an accountable officer where exempt in question C2.2.	UK only In EU Outside of EU
UK Digital Health Framework ¹⁰	Data Protection	NA	Please confirm where you store and process data (including any third party products your product uses)	Free text
UK Digital Health Framework ¹⁰	Data Protection	NA	If you process store or process data outside of the UK, please name	Free text

Framework Source	Category	Category Description	Question	Response Options
			the country and set out how the arrangements are compliant with current legislation	
UK Digital Health Framework ¹⁰	Technical Security	NA	Please attach your Cyber Essentials Certificate	Attached No evidence available
UK Digital Health Framework ¹⁰	Technical Security	NA	Please provide the summary report of an external penetration test of the product that included Open Web Application Security Project (OWASP) Top 10 vulnerabilities from within the previous 12 month period.	Attached No evidence available
UK Digital Health Framework ¹⁰	Technical Security		Please confirm whether all custom code had a security review.	Yes - Internal code review Yes - External code review No No because there is no custom code
UK Digital Health Framework ¹⁰	Technical Security		Please confirm whether all privileged accounts have appropriate Multi-Factor Authentication (MFA)?	Yes No
UK Digital Health Framework ¹⁰	Technical Security		Please confirm whether logging and reporting requirements have been clearly defined.	Yes No
UK Digital Health Framework ¹⁰	Technical Security		Please confirm whether the product has been load tested	Yes No
UK Digital Health Framework ¹⁰	Interoperability criteria		Does your product expose any Application Programme Interfaces (API) or integration channels for other consumers?	Yes No
UK Digital Health Framework ¹⁰	Interoperability criteria		If yes, please provide detail and evidence: <ul style="list-style-type: none"> • The API's (e.g. what they connect to) • Set out the healthcare standards of data interoperability eg. Health Level Seven International (HL7) / Fast Healthcare Interoperability Resources (FHIR) • Confirm that they follow Government Digital Services Open API Best Practice 	Free text

Framework Source	Category	Category Description	Question	Response Options
			<ul style="list-style-type: none"> Confirm they are documented and freely available Third parties have reasonable access to connect <p>If no, please set out why your product does not have APIs.</p>	
UK Digital Health Framework ¹⁰	Interoperability criteria		Do you use NHS number to identify patient record data?	Yes No No because product doesn't identify patient record data
UK Digital Health Framework ¹⁰	Interoperability criteria	NA	<p>If yes, please confirm whether it uses NHS Login to establish a user's verified NHS number.</p> <p>If no, please set out the rationale, how your product established NHS number and the associated security measures in place.</p>	Free text
UK Digital Health Framework ¹⁰	Interoperability criteria	NA	Does your product have the capability for read/write operations with electronic health records (EHRs) using industry standards for secure interoperability (e.g. OAuth 2.0, TLS 1.2)	Yes No No because the product doesn't read/write into EHRs
UK Digital Health Framework ¹⁰	Interoperability criteria	NA	If yes, please detail the standard	Free text
UK Digital Health Framework ¹⁰	Interoperability criteria	NA	If no, please state the reasons and mitigations, methodology and security measures.	Free text
UK Digital Health Framework ¹⁰	Interoperability criteria	NA	Is your product a wearable or device, or does it integrate with them?	Yes No
UK Digital Health Framework ¹⁰	Interoperability criteria	NA	If yes, provide evidence of how it complies with ISO/IEEE 11073 Personal Health Data (PHD) Standards.	Attached No evidence available
UK Digital Health Framework ¹⁰	Usability and accessibility - scored section	NA	Understand users and their needs in context of health and social care. Do you engage users in the development of the	Yes No Working towards it

Framework Source	Category	Category Description	Question	Response Options
			product?	
UK Digital Health Framework ¹⁰	Usability and accessibility - scored section	NA	If yes or working towards it, how frequently do you consider user needs in your product development and what methods do you use to engage users and understand their needs?	Free text
UK Digital Health Framework ¹⁰	Usability and accessibility - scored section	NA	Work towards solving a whole problem for users - Are all key user journeys mapped to ensure that the whole user problem is solved or it is clear to users how it fits into their pathway or journey?	Yes No Working towards it
UK Digital Health Framework ¹⁰	Usability and accessibility - scored section	NA	If yes or working towards it, please attach the user journeys and/or how the product fits into a user pathway or journey	Attached
UK Digital Health Framework ¹⁰	Usability and accessibility - scored section	NA	Make the service simple to use - Do you undertake user acceptance testing to validate usability of the system?	Yes No Working towards it
UK Digital Health Framework ¹⁰	Usability and accessibility - scored section	NA	If yes or working towards it, please attach information that demonstrates that user acceptance testing is in place to validate usability.	Attached
UK Digital Health Framework ¹⁰	Usability and accessibility - scored section	NA	Make sure everyone can use the service - Are you international Web Content Accessibility Guidelines (WCAG) 2.1 level AA compliant?	Yes No Working towards it
UK Digital Health Framework ¹⁰	Usability and accessibility - scored section	NA	Provide a link to your published accessibility statement.	Free text
UK Digital Health Framework ¹⁰	Usability and accessibility - scored section	NA	Create a team that includes multi-disciplinary skills and perspectives Does your team contain multidisciplinary skills?	Yes No Working towards it

Framework Source	Category	Category Description	Question	Response Options
UK Digital Health Framework ¹⁰	Usability and accessibility - scored section	NA	Use agile ways of working	Yes No Working towards it
UK Digital Health Framework ¹⁰	Usability and accessibility - scored section	NA	Do you use agile ways of working to deliver your product?	Yes No Working towards it
UK Digital Health Framework ¹⁰	Usability and accessibility - scored section	NA	Iterate and improve frequently Do you continuously develop your product?	Yes No Working towards it
UK Digital Health Framework ¹⁰	Usability and accessibility - scored section	NA	Define what success looks like and be open about how your service is performing - Do you have a benefits case that includes your objectives and the benefits you will be measuring and have metrics that you are tracking?	Yes No No because it isn't applicable
UK Digital Health Framework ¹⁰	Usability and accessibility - scored section	NA	Choose the right tools and technology - Does this product meet with NHS Cloud First Strategy? Does this product meet the NHS Internet First Policy?	Yes No No because it isn't applicable
UK Digital Health Framework ¹⁰	Usability and accessibility - scored section	NA	Use and contribute to open standards, common components and patterns - Are common components and patterns in use? If yes, which common components and patterns have been used?	Yes No Working towards it
UK Digital Health Framework ¹⁰	Usability and accessibility - scored section	NA	Operate a reliable service - Do you provide a Service Level Agreement to all customers purchasing the product?	Yes No
UK Digital Health Framework ¹⁰	Usability and accessibility - scored section	NA	Do you report to customers on your performance with respect to support, system performance (response times) and availability (uptime) at a frequency required by your customers? Please attach a copy of the information	Yes No

Framework Source	Category	Category Description	Question	Response Options
			provided to customers	
UK Digital Health Framework ¹⁰	Usability and accessibility - scored section	NA	Please attach a copy of the information provided to customers	Attached No evidence available
UK Digital Health Framework ¹⁰	Usability and accessibility - scored section	NA	Please provide your average service availability for the past 12 months, as a percentage to two decimal places	Free text

NA =not available

References for Appendix B

1. Stoyanov SR, Hides L, Kavanagh DJ, et al. Development and Validation of the User Version of the Mobile Application Rating Scale (uMARS). *JMIR Mhealth Uhealth*. 2016 Jun 10;4(2):e72. doi: 10.2196/mhealth.5849. PMID: 27287964.
2. Lagan S, Sandler L, Torous J. Evaluating evaluation frameworks: a scoping review of frameworks for assessing health apps. *BMJ Open*. 2021 Mar 19;11(3):e047001. doi: 10.1136/bmjopen-2020-047001. PMID: 33741674.
3. New Zealand health information | Health Navigator NZ. Accessed May 11, 2021. <https://www.healthnavigator.org.nz/>.
4. APA App Advisor. Accessed May 11, 2021. <https://www.psychiatry.org/psychiatrists/practice/mental-health-apps>.
5. App Guide | PsyberGuide. One Mind PsyberGuide. Accessed May 11, 2021. <https://onemindpsyberguide.org/apps/>.
6. Martinengo L, Van Galen L, Lum E, et al. Suicide prevention and depression apps' suicide risk assessment and management: a systematic assessment of adherence to clinical guidelines. *BMC Med*. 2019 Dec 19;17(1):231. doi: 10.1186/s12916-019-1461-z. PMID: 31852455.
7. Moshi MR, Tooher R, Merlin T. SUITABILITY OF CURRENT EVALUATION FRAMEWORKS FOR USE IN THE HEALTH TECHNOLOGY ASSESSMENT OF MOBILE MEDICAL APPLICATIONS: A SYSTEMATIC REVIEW. *Int J Technol Assess Health Care*. 2018 Jan;34(5):464-75. doi: 10.1017/s026646231800051x. PMID: 30201060.
8. Webb TL, Joseph J, Yardley L, et al. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *J Med Internet Res*. 2010 Feb 17;12(1):e4. doi: 10.2196/jmir.1376. PMID: 20164043.
9. Ondersma SJ, Scott TW. Clinician's Guide to Evaluating and Developing eHealth Interventions for Mental Health. *Psychiatric Research and Clinical Practice*. 2020;2(1):26-33. doi: 10.1176/appi.prep.2020.20190036.
10. Digital Technology Assessment Criteria (DTAC). NHSX. Accessed May 11, 2021. <https://www.nhsx.nhs.uk/key-tools-and-info/digital-technology-assessment-criteria-dtac/>.

Appendix C. Modifications in uMARS Questions

Table C-1. Modifications in uMARS questions

uMARS Question Topic - Question Number	Question	Type of Modification	uMARS Language	Our Section-Question Number	Our Framework Language
Quality -13	Quality of information: Is the app content correct, well written, and relevant to the goal and/or topic of the app?	Clarify question	Quality of information: Is the app content correct, well written, and relevant to the goal and/or topic of the app?"	2I -40	Is the app copy, well written, and relevant to the goal and/or topic of the app?
Quality -13	Quality of information: Is the app content, well written, and relevant to the goal and/or topic of the app?	Clarify response	Poor. Barely relevant/appropriate/coherent/may be incorrect	2I -40	Barely relevant/appropriate/coherent/may be incorrect"
Quality-13	Quality of information: Is the app content, well written, and relevant to the goal and/or topic of the app?	Clarify response	Moderately relevant/appropriate/coherent/and appears correct	2I -40	Satisfactory with respect to relevance/appropriateness/coherence and appears to be correct
Target group -5	Target group: Is the app content (visuals, language, design) appropriate for the target audience?	Clarify response	Acceptable but not specifically designed for the target audience, may be inappropriate/unclear/confusing at times	2I -31	Satisfactory but not specifically designed for the target audience, may be inappropriate/unclear/confusing at times
Graphics -11	Graphics: How high is the quality/resolution of graphics used for buttons, icons, menus, and content?	Clarify response	Moderate quality graphics and visual design, generally consistent in style	2I -33	Satisfactory quality graphics and visual design, generally consistent in style
Visual appeal -12	Visual appeal: How good does the app look?	Clarify response	OK: average, neither pleasant nor unpleasant	2I -34	Satisfactory: average, neither pleasant nor unpleasant
Gestural design -9	Gestural design: Do taps/swipes/pinches /scrolls make sense? Are they consistent across all components/screens?	Clarify response	OK with some inconsistencies/confusing elements	2I -39	Satisfactory with some inconsistencies/confusing elements

Appendix D. Pilot Rounds Results

Modifications to the Framework by Pilot Round

Table D-1. Modifications between pre-pilot round and pilot round 1

Section	Categories (Total Number of Questions)	Modified the Question Language for Clarity	Removed Question	Added Question To Further Capture the Concept	Added Additional Guidance Notes	Added or Consolidated Options for Responses	Added Question That Is Not Asked Elsewhere	Question Rearranged Sequentially in Framework	Total Modifications Per Category
SECTION 1	App integrity (4)	2	1	0	2	4	0	0	9
	Risk assessment (5)	2	3	0	1	2	1	0	9
	Evidence (3)	1	0	0	1	1	0	1	4
	Linkage to care (2)	0	0	0	1	1	1	1	4
	Access to crisis services (2)	1	1	0	0	0	0	1	3
	Vulnerable populations (0)	0	0	0	1	1	0	0	2
SECTION 2	Accessibility features (3)	0	0	0	0	3	0	0	3
	App information (2)	0	1	0	0	0	0	0	1
	Costs (3)	2	1	0	0	2	1	0	6
	Organizational credibility (3)	1	1	0	0	2	0	0	4
	Evidence and clinical foundation (5)	0	2	1	0	2	0	0	5
	Privacy and security (6)	1	1	1	0	5	2	0	10
	Informed consent (2)	0	0	0	0	2	0	0	2
	Cultural competence (2)	1	0	0	0	2	3	0	6
	Usability (13)	1	0	0	0	0	0	13	14
	Remote monitoring (4)	1	0	0	0	0	0	4	5
	Access to crisis response (1)	0	0	0	0	1	0	0	1
	Artificial Intelligence (1)	1	1	0	0	1	0	0	3
SECTION 3*	App features	NE	NE	NE	NE	NE	NE	NE	NE
	Functionality supported by the app (15)	NE	NE	NE	NE	NE	NE	NE	NE

This table only includes modifications made to questions that were used for interrater reliability testing ("base" questions). There may be additional modifications that were made to Pre- and Post-Assessment questions and other non-base questions that are not included in this tabulation.

* Section 3 was not evaluated in the pilot round 1 as it was not ready

NE=not evaluated

Table D-2. Modifications of framework between pilot round 1 and pilot round 2

Section	Categories (Total Number of Questions)	Modified the Question Language for Clarity	Removed Question	Added Question To Further Capture the Concept	Added Additional Guidance Notes	Added or Consolidated Options for Responses	Added Question That Is Not Asked Elsewhere	Question Rearranged Sequentially in Framework	Total Modifications Per Category
SECTION 1	App integrity (4)	1	0	0	4	0	1	0	6
	Risk assessment (3)	1	0	1	1	0	0	0	3
	Evidence (2)	0	1	0	1	0	0	0	2
	Linkage to care (1)	0	0	0	0	0	0	0	0
	Access to crisis services (1)	0	0	2	0	2	0	0	4
	Vulnerable populations (2)	0	1	0	0	0	1	0	2
SECTION 2	Accessibility features (3)	2	1	0	1	0	0	0	4
	App information (2)	0	0	0	1	0	0	0	1
	Costs (3)	0	0	0	1	2	0	0	3
	Organizational credibility (2)	0	0	0	0	1	0	0	1
	Evidence and clinical foundation (1)	0	0	0	0	0	0	0	0
	Privacy and security (8)	1	0	4	0	0	0	3	8
	Informed consent (1)	0	0	0	0	0	0	0	0
	Cultural competence (5)	1	0	2	2	1	0	0	6
	Usability (14)	0	0	0	0	0	0	0	0
	Content and duration of use (3)	0	0	0	0	0	2	1	3
	Remote monitoring (4)	0	0	0	0	4	0	0	4
	Access to crisis response (1)	0	0	0	0	1	0	0	1
	Artificial Intelligence (1)	0	0	0	0	0	0	0	0
SECTION 3 *	App features	NE	NE	NE	NE	NE	NE	NE	NE
	Functionality supported by the app	NE	NE	NE	NE	NE	NE	NE	NE

This table only includes modifications made to questions that were used for interrater reliability testing ("base" questions). There may be additional modifications that were made to Pre- and Post-Assessment questions and other non-base questions that are not included in this tabulation.

* Section 3 was not evaluated in the pilot round 1 as it was not ready

NE=not evaluated

Table D-3. Modifications between pilot round 2 and pilot round 3

Section	Categories (Total Number of Questions)	Modified the Question Language for Clarity	Removed Question	Added Question To Further Capture the Concept	Added Additional Guidance Notes	Added or Consolidated Options for Responses	Added Question That Is Not Asked Elsewhere	Question Rearranged Sequentially in Framework	Total Modifications Per Category
SECTION 1	App integrity (4)	0	0	0	0	0	0	0	0
	App Integrity Assessment (1)	0	0	0	1	0	0	0	1
	Risk (3)	0	0	0	0	1	1	0	2
	Risk Assessment (1)	0	0	0	0	0	0	0	0
	Evidence (2)	0	0	0	0	0	0	0	0
	Linkage to care (1)	0	0	0	0	0	0	0	0
	Access to crisis services (1)	0	0	0	0	1	0	0	1
SECTION 2	Vulnerable populations (2)	0	0	0	0	0	0	0	0
	Accessibility features (2)	0	0	0	2	1	0	0	3
	App information (2)	0	0	0	0	0	0	0	0
	Costs (3)	0	0	0	0	1	0	0	1
	Organizational credibility (2)	0	0	0	1	1	0	0	2
	Evidence and clinical foundation (1)	0	0	0	0	0	0	0	0
	Privacy and security (7)	0	0	0	2	2	0	0	4
	Informed consent (1)	0	0	0	0	1	0	0	1
	Cultural competence (5)	0	0	0	0	0	0	0	0
	Usability (13)	0	0	2	0	0	0	0	2
	Content and duration of use (3)	0	0	0	0	0	0	0	0
	Remote monitoring (4)	0	0	0	1	0	0	0	1
	Access to crisis response (1)	0	0	0	0	0	0	1	1
	Artificial Intelligence (1)	0	0	0	0	0	0	0	0
SECTION 3	App features (8)	0	0	0	0	0	0	0	0
	Functionality supported by the app (14)	1	1	1	1	0	1	0	5

This table only includes modifications made to questions that were used for interrater reliability testing ("base" questions). There may be additional modifications that were made to Pre- and Post-Assessment questions and other non-base questions that are not included in this tabulation.

Table D-4. Modifications between pilot round 3 and pilot round 4

Section	Categories (Total Number of Questions)	Modified the Question Language for Clarity	Removed Question	Added Question To Further Capture the Concept	Added Additional Guidance Notes	Added or Consolidated Options for Responses	Added Question That Is Not Asked Elsewhere	Question Rearranged Sequentially in Framework	Total Modifications Per Category
SECTION 1	App integrity (4)	0	0	0	0	0	0	0	0
	App Integrity Assessment (1)	0	0	0	0	0	0	0	0
	Risk (3)	0	0	0	0	0	0	0	0
	Risk Assessment (1)	0	0	0	0	0	0	0	0
	Evidence (2)	0	0	0	0	0	0	0	0
	Linkage to care (1)	1	0	0	0	0	0	0	1
	Access to crisis services (1)	0	0	0	0	0	0	0	0
	Risk Level (1)	0	0	0	0	0	0	0	0
	Vulnerable Populations (2)	0	0	0	0	0	0	0	0
SECTION 2	Accessibility features (2)	0	0	0	0	0	0	0	0
	App information (2)	0	0	0	0	0	0	0	0
	Costs (3)	0	0	0	0	0	0	0	0
	Organizational Credibility (2)	0	0	0	1	0	0	0	1
	Evidence and Clinical Foundation (1)	0	0	0	0	0	0	0	0
	Privacy and Security (7)	0	0	0	0	0	0	0	0
	Informed Consent (1)	0	0	0	0	0	0	0	0
	Cultural competence (5)	0	0	0	0	0	0	0	0
	Usability (14)	0	0	0	0	2	0	0	2
	Content and Duration of Use (3)	0	0	0	0	0	0	0	0
	Remote Monitoring (4)	0	0	0	0	0	0	0	0
	Access to Crisis Response (1)	0	0	0	0	0	0	0	0
	Artificial Intelligence (1)	0	0	0	0	0	0	0	0
SECTION 3	App features (7)	0	0	0	0	0	0	0	0
	Functionality supported by the app (15)	0	0	0	0	1	0	0	1

This table only includes modifications made to questions that were used for interrater reliability testing ("base" questions). There may be additional modifications that were made to Pre- and Post-Assessment questions and other non-base questions that are not included in this tabulation.

Table D-5. Modifications between pilot round 4 and pilot round 5

Section	Categories (Total Number of Questions)	Modified the Question Language for Clarity	Removed Question	Added Question To Further Capture the Concept	Added Additional Guidance Notes	Added or Consolidated Options for Responses	Added Question That Is Not Asked Elsewhere	Question Rearranged Sequentially in Framework	Total Modifications Per Category
SECTION 1	App integrity (4)	0	0	0	1	1	0	0	2
	App Integrity Assessment (1)	0	0	0	0	0	0	0	0
	Risk (1)	0	0	0	0	0	0	0	0
	Risk Assessment (1)	0	0	0	0	0	0	0	0
	Evidence (2)	0	0	0	0	0	0	0	0
	Linkage to care (1)	1	0	0	0	0	0	0	1
	Access to crisis services (1)	0	0	0	0	0	0	0	0
	Risk Level (1)	0	0	0	0	0	0	0	0
SECTION 2	Vulnerable Populations (2)	0	0	0	1	0	0	0	1
	Accessibility features (2)	0	0	0	0	0	0	0	0
	App information (2)	0	0	0	0	0	0	0	0
	Costs (4)	0	0	1	0	1	0	0	2
	Organizational Credibility (2)	0	0	0	1	0	0	0	1
	Evidence and Clinical Foundation (1)	0	0	0	0	0	0	0	0
	Privacy and Security (8)	0	0	0	0	1	0	1	2
	Informed Consent (1)	0	0	1	0	0	0	0	1
	Cultural competence (5)	0	0	0	0	0	0	0	0
	Usability (13)	0	0	0	0	1	0	0	1
	Content and Duration of Use (2)	0	0	0	0	0	0	0	0
	Remote Monitoring (3)	0	0	0	0	0	0	0	0
	Access to Crisis Response (1)	0	0	0	0	0	0	0	0
Artificial Intelligence (1)	0	0	0	0	0	0	0	0	
SECTION 3	App features (7)	0	0	0	0	0	0	0	0
	Functionality supported by the app (15)	0	0	0	0	1	0	0	1

This table only includes modifications made to questions that were used for interrater reliability testing ("base" questions). There may be additional modifications that were made to Pre- and Post-Assessment questions and other non-base questions that are not included in this tabulation.

Table D-6. Modifications between pilot round 5 and pilot round 6

Section	Categories (Total Number of Questions)	Modified the Question Language for Clarity	Removed Question	Added Question To Further Capture the Concept	Added Additional Guidance Notes	Added or Consolidated Options for Responses	Added Question That Is Not Asked Elsewhere	Question Rearranged Sequentially in Framework	Total Modifications Per Category
SECTION 1	App integrity (4)	0	0	0	0	0	0	0	0
	App Integrity Assessment (1)	0	0	0	0	0	0	0	0
	Risk (1)	0	0	0	0	0	0	0	0
	Risk Assessment (1)	0	0	0	0	0	0	0	0
	Evidence (2)	0	0	0	0	1	0	0	1
	Linkage to care (1)	0	0	0	0	1	0	1	2
	Access to crisis services (1)	0	0	0	0	0	0	0	0
	Risk Level (1)	0	0	0	0	0	0	0	0
SECTION 2	Vulnerable Populations (2)	0	0	0	1	0	0	0	1
	Accessibility features (2)	0	0	0	0	0	0	0	0
	App information (2)	0	0	0	0	0	0	0	0
	Costs (4)	0	0	0	1	1	0	0	2
	Organizational Credibility (2)	0	0	0	0	0	0	0	0
	Evidence and Clinical Foundation (1)	0	0	0	0	1	0	0	1
	Privacy and Security (6)	0	3	0	1	5	0	1	10
	Informed Consent (1)	0	0	0	1	0	0	0	1
	Cultural competence (5)	0	0	0	0	1	0	0	1
	Usability	0	0	0	0	0	0	0	0
	Content and Duration of Use (2)	0	0	0	0	0	0	0	0
	Remote Monitoring (3)	0	0	0	0	2	0	0	2
	Access to Crisis Response (1)	0	0	0	0	0	0	0	0
Artificial intelligence (1)	0	0	0	1	1	0	0	2	
SECTION 3	App features (7)	0	0	0	0	0	0	0	0
	Functionality supported by the app (15)	0	0	0	0	0	0	0	0

This table only includes modifications made to questions that were used for interrater reliability testing ("base" questions). There may be additional modifications that were made to Pre- and Post-Assessment questions and other non-base questions that are not included in this tabulation.

Interrater Agreement Results by Pilot Round

Table D-7. Interrater reliability for the pre-pilot round

App	Interrater Reliability – % Agreement
Breathe, Think, Do	63.6%
BUP	72.7%
Daylio	75.0%
Replika	81.3%
Sleepio	81.8%
SuperBetter	87.5%
T2 Mood Tracker	62.5%
The Safe Place	81.8%
Woebot	62.5%
Youper	45.5%

Table D-8. Interrater agreement result for the pilot round 1

Section Categories	N	Type of Summarization	App1 Narcolepsy Monitor	App2 WeAreMore	App 3 Awell	App 4 Pursue Care	App 5 MDLIVE	Overall Agreement Across Apps for Round 1
SECTION 1 App integrity	4	% items within each category for which there was interrater agreement	100.00%	75.00%	75.00%	25.00%	75.00%	70.0%
SECTION 1 Risk assessment	3	Agreement/Disagreement on level of risk	66.67%	66.67%	66.67%	33.33%	33.33%	53.3%
SECTION 1 Evidence	2	% items within each category for which there was interrater agreement	50.00%	100.00%	0.00%	100.00%	50.00%	60.0%
SECTION 1 Linkage to care	1	% items within each category for which there was interrater agreement	100.00%	100.00%	0.00%	100.00%	50.00%	70.0%
SECTION 1 Access to crisis services	1	% items within each category for which there was interrater agreement	100.00%	100.00%	0.00%	0.00%	0.00%	40.0%
SECTION 1 Vulnerable populations	2	% items within each category for which there was interrater agreement	50.00%	50.00%	50.00%	50.00%	50.00%	50.0%
SECTION 2 Accessibility features	3	% items within each category for which there was interrater agreement	100.00%	66.67%	33.33%	100.00%	33.33%	66.7%
SECTION 2 App information	2	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	100.0%
SECTION 2 Costs	3	% items within each category for which there was interrater agreement	66.67%	33.33%	33.33%	33.33%	66.67%	46.7%
Organizational Credibility	2	% items within each category for which there was interrater agreement	50.00%	100.00%	100.00%	50.00%	50.00%	70.0%
SECTION 2 Evidence and Clinical Foundation	1	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	100.0%
SECTION 2 Privacy and Security	8	% items within each category for which there was interrater agreement	62.5%	12.5%	75.0%	87.5%	25.0%	52.5%

Section Categories	N	Type of Summarization	App1 Narcolepsy Monitor	App2 WeAreMore	App 3 Awell	App 4 Pursue Care	App 5 MDLIVE	Overall Agreement Across Apps for Round 1
SECTION 2 Informed Consent	1	% items within each category for which there was interrater agreement	100.0%	100.0%	100.0%	0.0%	100.0%	80.0%
SECTION 2 Cultural competence	5	% items within each category for which there was interrater agreement	60.0%	60.0%	60.0%	100.0%	100.0%	76.0%
SECTION 2 Usability	14	Summarize by condensing responses on each end	85.7%	92.9%	92.9%	64.3%	85.7%	84.3%
SECTION 2 Content and Duration of Use	3	% items within each category for which there was interrater agreement	0.0%	66.7%	33.3%	33.3%	100.0%	46.7%
SECTION 2 Remote Monitoring	4	% items within each category for which there was interrater agreement	100.0%	75.0%	50.0%	50.00%	50.0%	65.0%
SECTION 2 Access to Crisis Response	1	% items within each category for which there was interrater agreement	100.0%	100.0%	100.0%	100.0%	0.0%	80.0%
SECTION 2 Artificial Intelligence	1	% items within each category for which there was interrater agreement	100.0%	100.0%	100.0%	100.0%	0.0%	80.0%
SECTION 3* App features	NE	NE	NE	NE	NE	NE	NE	NE
SECTION 3* Functionality supported by the app	NE	NE	NE	NE	NE	NE	NE	NE

*Section 3 was not ready for the pilot round 1.

NE = not evaluated; N = total number of items

Table D-9. Interrater agreement result for the pilot round 2

Section	N	Type of Summarization	App1 StressScan	App2 Mental Health Tests - Mind Diagnostics	App 3 Change Scope	App 4 Diet or Disorder	App 5 FearTools	Overall Agreement Across Apps for Round 2
SECTION 1 App integrity	4	% items within each category for which there was interrater agreement	75.00%	100.00%	100.00%	50.00%	100.00%	85.00%
SECTION 1 App Integrity Assessment	1	Agreement/Disagreement on level of risk	0.00%	100.00%	100.00%	0.00%	100.00%	60.00%
SECTION 1 Risk	3	% items within each category for which there was interrater agreement	100.00%	66.67%	66.67%	66.67%	100.00%	80.00%
SECTION 1 Risk Assessment	1	Agreement/Disagreement on level of risk	100.00%	0.00%	0.00%	0.00%	100.00%	40.00%
SECTION 1 Evidence	2	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	50.00%	50.00%	80.00%
SECTION 1 Linkage to care	1	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 1 Access to crisis services	1	% items within each category for which there was interrater agreement	100.00%	100.00%	0.00%	100.00%	100.00%	80.00%
SECTION 1 Vulnerable Populations	2	% items within each category for which there was interrater agreement	100.00%	50.00%	0.00%	0.00%	100.00%	50.00%
SECTION 2 Accessibility features	2	% items within each category for which there was interrater agreement	50.00%	50.00%	0.00%	50.00%	50.00%	40.00%
SECTION 2 App information	2	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

Section Categories	N	Type of Summarization	App1 StressScan	App2 Mental Health Tests - Mind Diagnostics	App 3 Change Scope	App 4 Diet or Disorder	App 5 FearTools	Overall Agreement Across Apps for Round 2
SECTION 2 Costs	3	% items within each category for which there was interrater agreement	66.67%	100.00%	100.00%	100.00%	100.00%	93.33%
SECTION 2 Organizational Credibility	2	% items within each category for which there was interrater agreement	100.00%	50.00%	100.00%	100.00%	50.00%	80.00%
SECTION 2 Evidence and Clinical Foundation	1	% items within each category for which there was interrater agreement	0.00%	100.00%	0.00%	0.00%	100.00%	40.00%
SECTION 2 Privacy and Security	7	% items within each category for which there was interrater agreement	57.14%	42.86%	42.86%	100%	100.00%	68.57%
SECTION 2 Informed Consent	1	% items within each category for which there was interrater agreement	100.00%	100.00%	0.00%	0.00%	0.00%	40.00%
SECTION 2 Cultural competence	5	% items within each category for which there was interrater agreement	80.00%	80.00%	80.00%	100.00%	80.00%	84.00%
SECTION 2 Usability	13	Summarize by condensing responses on each end	76.92%	76.92%	76.92%	38.46%	61.54%	66.15%
SECTION 2 Content and Duration of Use	3	% items within each category for which there was interrater agreement	66.67%	66.67%	66.67%	33.33%	66.67%	60.00%
SECTION 2 Remote Monitoring	4	% items within each category for which there was interrater agreement	83.33%	83.33%	83.33%	100.00%	100.00%	90.00%

Section Categories	N	Type of Summarization	App1 StressScan	App2 Mental Health Tests - Mind Diagnostics	App 3 Change Scope	App 4 Diet or Disorder	App 5 FearTools	Overall Agreement Across Apps for Round 2
SECTION 2 Access to Crisis Response	1	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 2 Artificial Intelligence	1	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 3 App features	8	% items within each category for which there was interrater agreement	100.00%	75.00%	87.50%	100.00%	100.00%	92.50%
SECTION 3 Functionality supported by the app	14	Summarize by condensing responses on each end	85.71%	78.57%	92.86%	71.43%	57.14%	77.14%

Table D-10. Interrater agreement result for the pilot round 3

Section	N	Type of Summarization	App1 Self-Harm Recovery	App2 Schizophrenia HealthStorylines	App 3 Stanley-Brown Safety Plan	App 4 PsychSurveys	Overall Agreement Across Apps for Round 3
SECTION 1 App integrity	4	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	50.00%	87.50%
SECTION 1 App Integrity Assessment	1	Agreement/Disagreement on level of risk	100.00%	100.00%	100.00%	0.00%	75.00%
SECTION 1 Risk	3	Agreement/Disagreement on level of risk	100.00%	100.00%	100.00%	66.67%	91.67%
SECTION 1 Risk Assessment	1	Agreement/Disagreement on level of risk	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 1 Evidence	2	% items within each category for which there was interrater agreement	100.00%	100.00%	50.00%	100.00%	87.50%
SECTION 1 Linkage to care	1	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 1 Access to crisis services	1	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	0.00%	75.00%
SECTION 1 Risk Level	1	Agreement/Disagreement on level of risk	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 1 Vulnerable Populations	2	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 2 Accessibility features	2	% items within each category for which there was interrater agreement	100.00%	50.00%	100.00%	0.00%	62.50%
SECTION 2 App information	2	% items within each category for which there was interrater agreement	50.00%	100.00%	50.00%	100.00%	75.00%

Section Categories	N	Type of Summarization	App1 Self-Harm Recovery	App2 Schizophrenia HealthStorylines	App 3 Stanley-Brown Safety Plan	App 4 PsychSurveys	Overall Agreement Across Apps for Round 3
SECTION 2 Costs	3	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 2 Organizational Credibility	2	% items within each category for which there was interrater agreement	50.00%	100.00%	100.00%	50.00%	75.00%
SECTION 2 Evidence and Clinical Foundation	1	% items within each category for which there was interrater agreement	0.00%	100.00%	100.00%	0.00%	50.00%
SECTION 2 Privacy and Security	7	% items within each category for which there was interrater agreement	87.50%	57.14%	71.43%	28.57%	61.16%
SECTION 2 Informed Consent	1	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 2 Cultural competence	5	% items within each category for which there was interrater agreement	100.00%	100.00%	80.00%	100.00%	95.00%
SECTION 2 Usability	14	Summarize by condensing responses on each end	64.29%	71.43%	85.71%	85.71%	76.79%
SECTION 2 Content and Duration of Use	3	% items within each category for which there was interrater agreement	66.67%	100.00%	100.00%	66.67%	83.33%
SECTION 2 Remote Monitoring	4	% items within each category for which there was interrater agreement	100.00%	66.67%	66.67%	83.33%	79.17%
SECTION 2 Access to Crisis Response	1	% items within each category for which there was interrater agreement	0.00%	100.00%	100.00%	0.00%	50.00%
SECTION 2 Artificial Intelligence	1	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 3 App features	7	% items within each category for which there was interrater agreement	100.00%	85.71%	100.00%	85.71%	92.86%

Section Categories	N	Type of Summarization	App1 Self-Harm Recovery	App2 Schizophrenia HealthStorylines	App 3 Stanley-Brown Safety Plan	App 4 PsychSurveys	Overall Agreement Across Apps for Round 3
SECTION 3 Functionality supported by the app	15	Summarize by condensing responses on each end	93.33%	73.33%	80.00%	86.67%	83.33%

N=total number of items

Table D-11. Interrater agreement result for the pilot round 4

Section	N	Type of Summarization	App 1 Calm	App 2 Joyster	App 3 Simple Habit (Sleep, Meditation)	Overall Agreement Across Apps for Round 4
SECTION 1 App integrity	4	% items within each category for which there was interrater agreement	100.00%	75.00%	100.00%	91.67%
SECTION 1 App Integrity Assessment	1	Agreement/Disagreement on level of risk	100.00%	100.00%	100.00%	100.00%
SECTION 1 Risk	3	Agreement/Disagreement on level of risk	100.00%	100.00%	100.00%	100.00%
SECTION 1 Risk Assessment	1	Agreement/Disagreement on level of risk	100.00%	100.00%	100.00%	100.00%
SECTION 1 Evidence	2	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%
SECTION 1 Linkage to care	1	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%
SECTION 1 Access to crisis services	1	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%
SECTION 1 Risk Level	1	Agreement/Disagreement on level of risk	100.00%	100.00%	100.00%	100.00%
SECTION 1 Vulnerable Populations	2	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%
SECTION 2 Accessibility features	2	% items within each category for which there was interrater agreement	50.00%	100.00%	50.00%	66.67%
SECTION 2 App information	2	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%
SECTION 2 Costs	3	% items within each category for which there was interrater agreement	100.00%	66.67%	100.00%	88.89%

Section Categories	N	Type of Summarization	App 1 Calm	App 2 Joyster	App 3 Simple Habit (Sleep, Meditation)	Overall Agreement Across Apps for Round 4
SECTION 2 Organizational Credibility	2	% items within each category for which there was interrater agreement	100.00%	100.00%	50.00%	83.33%
SECTION 2 Evidence and Clinical Foundation	1	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%
SECTION 2 Privacy and Security	7	% items within each category for which there was interrater agreement	100.00%	71.43%	71.43%	80.95%
SECTION 2 Informed Consent	1	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%
SECTION 2 Cultural competence	5	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%
SECTION 2 Usability	13	Summarize by condensing responses on each end	92.31%	85.71%	85.71%	87.91%
SECTION 2 Content and Duration of Use	2	% items within each category for which there was interrater agreement	50.00%	100.00%	50.00%	66.67%
SECTION 2 Remote Monitoring	4	% items within each category for which there was interrater agreement	75.00%	75.00%	100.00%	83.33%
SECTION 2 Access to Crisis Response	1	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%
SECTION 2 Artificial Intelligence	1	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%
SECTION 3 App features	7	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%
SECTION 3 Functionality supported by the app	15	Summarize by condensing responses on each end	60.00%	100.00%	53.33%	71.11%

N= total number of items

Table D-12. Interrater agreement result for the fifth pilot round

Section Categories	N	Type of Summarization	App 1 Tourette Syndrome	App 2 Pattern: Correlate, Health Diary, Mood Tracker	App 3 Mental Health Test	App 4 Cognitive Stimulation Questions	App 5 Quizzilla	App 6 Njoy	App 7 Narcolepsy Disorder	App 8 ShutEye: Sleep Tracker	Overall Agreement Across Apps for Round 5
SECTION 1 App integrity	4	% items within each category for which there was interrater agreement	100.00%	50.00%	100.00%	100.00%	75.00%	100.00%	100.00%	75.00%	87.50%
SECTION 1 App Integrity Assessment	1	Agreement/Disagreement on level of risk	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 1 Risk	1	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 1 Risk Assessment	1	Agreement/Disagreement on level of risk	100.00%	100.00%	0.00%	100.00%	0.00%	100.00%	100.00%	0.00%	62.50%
SECTION 1 Evidence	2	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 1 Linkage to care	1	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 1 Access to crisis services	1	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 1 Risk Level	1	Agreement/Disagreement on level of risk	100.00%	100.00%	0.00%	100.00%	0.00%	100.00%	100.00%	0.00%	62.50%
SECTION 1 Vulnerable Populations	2	% items within each category for which there was interrater agreement	100.00%	100.00%	50.00%	100.00%	50.00%	100.00%	100.00%	50.00%	81.25%

Section Categories	N	Type of Summarization	App 1 Tourette Syndrome	App 2 Pattern: Correlate, Health Diary, Mood Tracker	App 3 Mental Health Test	App 4 Cognitive Stimulation Questions	App 5 Quitzilla	App 6 Njoy	App 7 Narcolepsy Disorder	App 8 ShutEye: Sleep Tracker	Overall Agreement Across Apps for Round 5
SECTION 2 Accessibility features	2	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 2 App information	2	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 2 Costs	4	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	75.00%	100.00%	100.00%	100.00%	96.88%
SECTION 2 Organizational Credibility	2	% items within each category for which there was interrater agreement	100.00%	50.00%	50.00%	50.00%	100.00%	50.00%	100.00%	100.00%	75.00%
SECTION 2 Evidence and Clinical Foundation	1	% items within each category for which there was interrater agreement	100.00%	0.00%	100.00%	0.00%	100.00%	0.00%	0.00%	100.00%	50.00%
SECTION 2 Privacy and Security	8	% items within each category for which there was interrater agreement	62.50%	50.00%	75.00%	37.50%	62.50%	62.50%	50.00%	62.50%	57.81%
SECTION 2 Informed Consent	1	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	0.00%	100.00%	0.00%	75.00%
SECTION 2 Cultural competence	5	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	80.00%	100.00%	97.50%
SECTION 2 Usability	13	Summarize by condensing responses on each end	76.92%	46.15%	69.23%	76.92%	53.85%	84.62%	76.92%	92.31%	72.12%

Section Categories	N	Type of Summarization	App 1 Tourette Syndrome	App 2 Pattern: Correlate, Health Diary, Mood Tracker	App 3 Mental Health Test	App 4 Cognitive Stimulation Questions	App 5 Quitzilla	App 6 Njoy	App 7 Narcolepsy Disorder	App 8 ShutEye: Sleep Tracker	Overall Agreement Across Apps for Round 5
SECTION 2 Content and Duration of Use	2	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	50.00%	100.00%	100.00%	93.75%
SECTION 2 Remote Monitoring	3	% items within each category for which there was interrater agreement	66.67%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	95.83%
SECTION 2 Access to Crisis Response	1	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 2 Artificial Intelligence	1	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 3 App features	7	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 3 Functionality supported by the app	15	Summarize by condensing responses on each end	100.00%	80.00%	93.33%	86.67%	86.67%	73.33%	93.33%	86.67%	87.50%

N=total number of items

Table D-13. Interrater agreement result for the pilot round 6

Section Categories	N	Type of Summarization	App 1 Bipolar Disorder	App 2 Mental Health Tracker: Disorders Test	App 3 Quirk CBT	App 4 Reflectly - Journal and AI Diary	App 5 Meditopia	App 6 Relaxing Music 2021	App 7 Nina App	App 8 Breeze: Mood Tracker, Ddiary	App 9 Waking Up: Guided Meditation	App 10 Peptalk Motivation	Overall Agreement Across Apps
SECTION 1 App integrity	4	% items within each category for which there was interrater agreement	100.00%	100.00%	75.00%	100.00%	100.00%	100.00%	100.00%	75.00%	75.00%	100.00%	93.75%
SECTION 1 App Integrity Assessment	1	Agreement/ Disagreement on level of risk	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	0.00%	100.00%	100.00%	87.50%
SECTION 1 Risk	1	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 1 Risk Assessment	1	Agreement/ Disagreement on level of risk	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	0.00%	100.00%	100.00%	100.00%	87.50%
SECTION 1 Evidence	2	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	50.00%	50.00%	100.00%	93.75%
SECTION 1 Linkage to care	1	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 1 Access to crisis services	1	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 1 Risk Level	1	Agreement/ Disagreement on level of risk	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	0.00%	100.00%	100.00%	100.00%	87.50%

Section Categories	N	Type of Summarization	App 1 Bipolar Disorder	App 2 Mental Health Tracker: Disorders Test	App 3 Quirk CBT	App 4 Reflectly - Journal and AI Diary	App 5 Meditopia	App 6 Relaxing Music 2021	App 7 Nina App	App 8 Breeze: Mood Tracker, Ddiary	App 9 Waking Up: Guided Meditation	App 10 Peptalk Motivation	Overall Agreement Across Apps
SECTION 1 Vulnerable Populations	2	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	50.00%	100.00%	100.00%	100.00%	93.75%
SECTION 2 Accessibility features	2	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	50.00%	100.00%	50.00%	100.00%	100.00%	100.00%	87.50%
SECTION 2 App information	2	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 2 Costs	4	% items within each category for which there was interrater agreement	100.00%	100.00%	75.00%	100.00%	50.00%	100.00%	100.00%	75.00%	100.00%	100.00%	87.50%
SECTION 2 Organizational Credibility	2	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 2 Evidence and Clinical Foundation	1	% items within each category for which there was interrater agreement	100.00%	0.00%	100.00%	100.00%	100.00%	0.00%	100.00%	100.00%	100.00%	100.00%	75.00%
SECTION 2 Privacy and Security	6	% items within each category for which there was interrater agreement	66.67%	83.33%	100.00%	66.67%	83.33%	83.33%	83.33%	83.33%	100.00%	66.67%	81.25%

Section Categories	N	Type of Summarization	App 1 Bipolar Disorder	App 2 Mental Health Tracker: Disorders Test	App 3 Quirk CBT	App 4 Reflectly - Journal and AI Diary	App 5 Meditopia	App 6 Relaxing Music 2021	App 7 Nina App	App 8 Breeze: Mood Tracker, Ddiary	App 9 Waking Up: Guided Meditation	App 10 Peptalk Motivation	Overall Agreement Across Apps
SECTION 2 Informed Consent	1	% items within each category for which there was interrater agreement	100.00%	100.00%	0.00%	0.00%	100.00%	100.00%	100.00%	0.00%	0.00%	0.00%	62.50%
SECTION 2 Cultural competence	5	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	80.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	97.50%
SECTION 2 Usability	13	Summarize by condensing responses on each end	92.31%	46.15%	84.62%	100.00%	84.62%	46.15%	92.31%	84.62%	84.62%	76.92%	78.85%
SECTION 2 Content and Duration of Use	2	% items within each category for which there was interrater agreement	100.00%	100.00%	50.00%	50.00%	50.00%	50.00%	50.00%	50.00%	100.00%	100.00%	62.50%
SECTION 2 Remote Monitoring	3	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	66.67%	100.00%	100.00%	100.00%	66.67%	100.00%	100.00%	91.67%
SECTION 2 Access to Crisis Response	1	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 2 Artificial Intelligence	1	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	0.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	87.50%

Section Categories	N	Type of Summarization	App 1 Bipolar Disorder	App 2 Mental Health Tracker: Disorders Test	App 3 Quirk CBT	App 4 Reflectly - Journal and AI Diary	App 5 Meditopia	App 6 Relaxing Music 2021	App 7 Nina App	App 8 Breeze: Mood Tracker, Ddiary	App 9 Waking Up: Guided Meditation	App 10 Peptalk Motivation	Overall Agreement Across Apps
SECTION 3 App features	7	% items within each category for which there was interrater agreement	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
SECTION 3 Functionality supported by the app	15	Summarize by condensing responses on each end	100.00%	100.00%	80.00%	80.00%	86.67%	80.00%	73.33%	46.67%	86.67%	100.00%	80.83%

Appendix E. Framework Training Guide: Framework to Assist Stakeholders in Technology Evaluation for Recovery (FASTER) to Mental Health and Wellness Framework Assessment

The goal of this assessment framework is to support agencies and individuals working in mental health, as well as users of mental health apps, in making informed decisions about using or recommending the use of particular apps. We expect that this framework will be applied by intermediary mental health agencies that have the capacity to train personnel to use this framework to evaluate mental health apps and insurance companies that might have an interest in reimbursing certain health apps. Evaluators who use this framework and accompanying training guide should be individuals with some background in technology and mental health, and not required to be experts in technology or mental health. The **results** of such assessments will be valuable to healthcare professionals as they prescribe apps to patients, and to patients/users/caregivers in search for a mental health support app. The framework might also guide app developers in the development of apps. The framework consists of three sections: Risks and Mitigation Strategies, Function, and Mental Health App Features (Figure E-1), and categories of questions in each of the sections (Figure E-2).

Section 1: Risks and Mitigation Strategies. This section assesses risks posed by apps, as evaluated by alignment of the goals of the app, the target audience, and the severity of the mental health condition with available evidence to support the approach, oversight, and linkage to care, and privacy and security protocols. In this section, we also ask questions to assess the credibility of the developer.

Apps that do not meet the threshold for risk will be flagged accordingly. However the reviewer should continue with the assessment.

Section 2: Function. This section focuses on more descriptive aspects of accessibility, usability, privacy/security, and other relevant information about the app. Individual criterion may be scored or unscored but will result in narrative summaries about how the app scores within each category. The questions in this section are related to accessibility features, basic app information, costs, usability, organizational credibility, evidence and clinical foundation, artificial intelligence (AI) functionality and remote monitoring (Figure E-2).

Section 3: Mental Health App Features. This section focuses on features related to mental health (Figure E-2).

Figure E-1. Framework to Assist Stakeholders in Technology Evaluation for Recovery (FASTER)

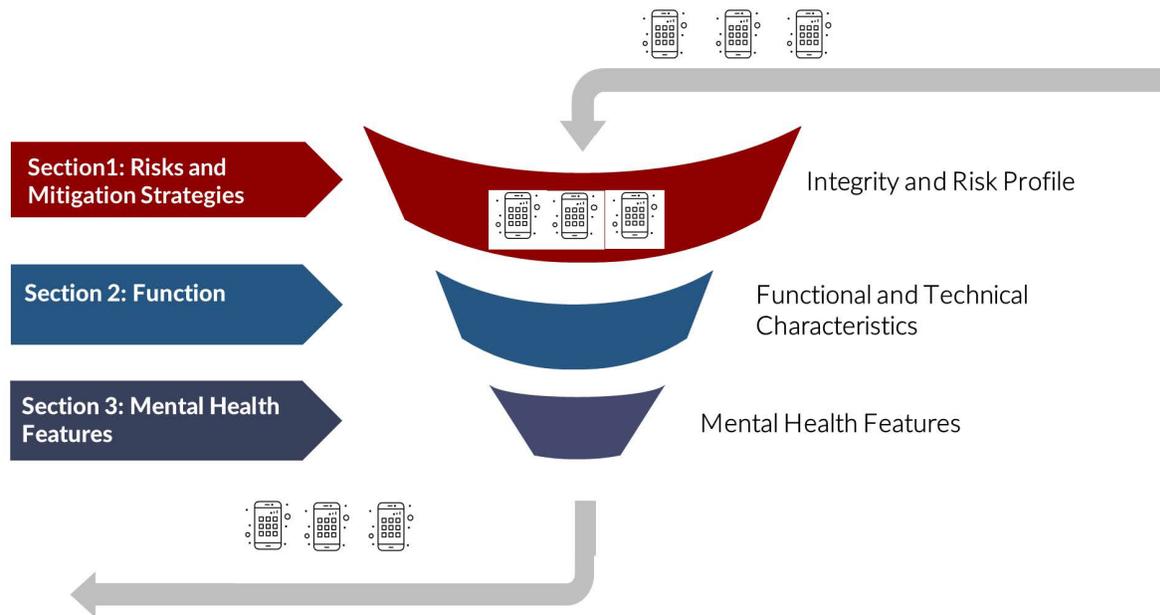


Figure E-2. Categories of questions in each section

	Risks and Mitigation Strategies	App Integrity	Risk Assessment	Evidence	Linkage to Care	Access to Crisis Resources
	Function	Accessibility Features	App Info	Costs	Organizational Credibility	Evidence & Clinical Foundation
		Privacy & Security	Informed Consent	Cultural Competence	Usability	Remote Monitoring
		Access to Crisis Resources	Artificial Intelligence (AI)			
	Mental Health Features	Mental Health Features (e.g., facilitating social interaction, motivation enhancement, planning/alternative strategies/planning for high -risk situations, screening, self -help, skill building, safety planning, and promoting sleep hygiene)				

In addition to these three sections, we have an administrative questionnaire at the beginning and the end of the assessment.

Logistics for App Download: Download the app from the App Store to your phone or tablet and create a login. Note the privacy and security information that is provided or requested when you create a login, because there are questions about privacy and security in the app screening process.

Apps That Require a Code: Some apps may require a code from an insurance provider or healthcare provider in order to create a login. In this case, please check the website for contact information for the organization, so that you can request access to the app for this screening purpose.

Searching for Evidence in PubMed: Combine the name of the app with the clinical query for *therapy*.

1. Go to PubMed Clinical: <https://pubmed.ncbi.nlm.nih.gov/clinical/>
2. Enter the **name of the app** in the search bar.
3. Query *therapy*.
4. Select ***broad*** under *Clinical Study Categories* on the right-hand side of the page.
5. Click *Search*.

Appendix F. Framework

The Framework to Assist Stakeholders in Technology Evaluation for Recovery (FASTER) to Mental Health and Wellness comprises three sections with an initial and concluding set of administrative questions: Section 1: Risks and Mitigation Strategies, Section 2: Function, and Section 3: Mental Health App Features. Within each of these sections, there are a series of items related to the assessment of specific categories considered critical.

Appendix F. Framework: Administrative Questionnaire

This section contains introductory questions about the app to be reviewed. The questions are mostly factual, rather than evaluative. Some of these questions may make it into the final Assessment Framework, based on the information we gather from the pilot.

Q 1. Reviewer Name

Q 2. Date of Evaluation

Q 3. App Name

Q 4. App/Company Website

This should only be a specific website developed by the app developer and not their pages on social media sites, such as Instagram or Facebook.

Q 5. Country of Origin

This information can be found in the Terms & Conditions on the website or in the app.

Q 6. App Version Number

This information can be found on the app and from the App Store.

Q 7. Platform Used for Assessment by the Reviewer

Q 8. OS Version Number on Evaluating Device

For Android go to: Settings > About Phone > Software Information;
For iOS go to: Settings > General > Software Version.

Q 9. Describe the goal of the app in your own words. Do not use any of the marketing phrases/lingo used by the app developer.

This should be based on the app description on the website.

Q 10. Has the app been approved by any regulating authority, such as FDA?

This information can be found on the app developer website.

Q 11. Does the app have multiple revenue models, such as freemium, in-app purchases, etc.?

- Yes
- No

Q 12. If the app has multiple versions (e.g., free, freemium, paid), which version did you evaluate? If the app had only one revenue model, please choose “not applicable” here.

- Free
- Paid version
- Paid by insurance agency, employer, or healthcare provider
- Not applicable
- Unable to assess

Q 13. Does the app require prior authorization from a healthcare institution/insurance provider/college/employer or other institution for access?

- Yes
- No
- Unable to assess

Q 14. If you were unable to assess the app, please specify the reason.

Note: Questions about privacy, security, and informed consent are covered in both Section 1 and Section 2 of the framework. Please make a note of all the disclaimers, warnings, and privacy and security statements and agreements when you first log into the app, so you are able to answer the questions that come later in the framework.

Section 1: Risks and Mitigation Strategies

Section 1 aims to evaluate the potential risk posed by the app, and evidence of efficacy, and safety features. The burden of evidence to prove efficacy and safety is dependent on the goals of the app, the target audience, and the severity of the mental health condition.

The categories of questions in this section are App Integrity, Risk Assessment, Evidence, Linkage to Care, and Access to Crisis Services. If apps do not meet the threshold for safety and credibility, that is flagged in the assessment.

A. App Integrity

The information requested in these questions can be found on the App Store, the app's website, or via Google Search.

1. Was the version of the app you are reviewing updated in the last 6 months?

- Yes
- No
- Unable to assess

2. Does the app provide a privacy and security agreement to be reviewed and agreed to by the user?

The agreement should be accessible through a link on the app or on the mobile app. Reviewers should not look for this information in other locations, such as the developer's website.

- Yes, there is a privacy and security agreement to be agreed to by the user
- No, there is not a privacy and security agreement
- Unable to assess or unable to access

3. Does the app provide warnings and disclaimers (e.g., the limitations of the app related to medical liability)?

Reviewers should look for this information on the mobile app ONLY.

- Yes, warnings and/or disclaimers are provided
- No, warnings and/or disclaimers are not provided
- Unable to assess
- Not applicable

4. Has the app been endorsed by, or is it being used by, a government agency or trusted mental health professional association?

Endorsement and/or usage helps us establish the credibility of the app.

Endorsement means to give one's approval to, especially officially by an organization. An app merely being reviewed by the American Psychological Association (APA), or a similar organization, does not mean it has been approved by them. Endorsement should be for the app and not the company developing the app or the domain (e.g., teletherapy). Self-reported endorsements are acceptable.

Usage by an institution signifies that an app is made available to all its members. For example, Veterans Affairs uses iBlueButton to deliver critical health care information to veterans.

Examples of government agencies include federal, state and local agencies such as CMS, Veterans Affairs, Virginia Department of Health. Examples of trusted mental health professional associations include Substance Abuse and Mental Health Services Administration (SAMHSA), American Psychological Association (APA), and National Alliance on Mental Illness (NAMI).

- Yes, the app has been endorsed by one or more mental health associations, government agencies, or non-government bodies (examples may include but are not limited to government agencies, such as Centers for Medicare & Medicaid Services (CMS) or Veterans Affairs (VA); non-government bodies, such as APA; and insurance or healthcare institutions)
- No, the app has not been endorsed
- Unable to assess, please specify reason: _____

App Integrity Assessment

Question	Response	App Integrity Level
Does the app satisfy the Integrity Assessment?	<input type="radio"/> Responses to questions 1-3 are all Yes OR response to question 4 is Yes	App Integrity is "High"
Please pick only 1 response.	<input type="radio"/> Response to any of the questions 1-3 is No AND response to 4 is No	App Integrity is "Low"

B. Risk Assessment

This category aims to determine the level of risk posed by the app. For example, if the app is used by someone with a clinically diagnosed condition, such as Schizophrenia, then the risk posed by the app may be higher than if the person has mild anxiety. Similarly, if the app targets children and adolescents, then the potential risk associated with the app may be higher than if it targeted adults with no impairment or mental health diagnosis.

*For the purposes of this report, either minors (younger than 18 years of age) or adults with a moderate or severe level of functional impairment that causes them to need support from a caregiver are considered to be part of a **vulnerable population**.*

The information requested in these questions can be found within the app/app website.

5. Is the app intended for use by adults who may have an illness or disability that impacts their decisionmaking ability? Severe autism or severe dementia are examples of illnesses or disabilities that may affect decisionmaking ability.

For the purposes of this assessment, the population that may have an illness or disability that impacts their decision making ability is included in the "vulnerable population" group. Adults can usually make medical decisions, such as consenting to treatment, on their own. In some cases, adults need someone else to make medical decisions for them if their decisionmaking ability is affected by an illness or disability. This is sometimes referred to as "substituted consent." A higher level of evidence/protection is needed for the use of specific apps under these circumstances.

- Yes
- No
- Unable to assess

6. Does the app claim that it is intended for use by minors (i.e., people younger than 18 years of age)?

For the purposes of this assessment, this population is included in the "vulnerable population" group. The app is "intended for use" by a population if it is clearly specified on the app or website that this population is its target user group. Do not look at the App Store Age Ratings or Terms and Conditions for this information.

- Yes, for use by minors
- Yes, for use by both minors and adults
- No, not for use by minors
- Unable to assess

7. Does the app claim to provide standalone treatment for any mental health condition?

A standalone treatment might be some form of psychotherapy or other any other medical care provided by the app independently, rather than in conjunction with a licensed healthcare practitioner.

Functionality, such as mindfulness training, skills training, symptom or mood tracking, and social support, would not be considered standalone treatment.

- Yes, the app provides standalone treatment
- No, the app does not provide standalone treatment

- Unable to assess

C. Evidence

It is important that the app has a solid clinical foundation. If the app targets a vulnerable population, the burden of evidence to prove efficacy and safety is greater.

The information requested in these questions can be found on the app website (look at customer testimonials or resources/blog for anecdotal information) and on PubMed (search PubMed as outlined in the training document). Make sure to save all the references in the app review document.

8. Has the app been evaluated for efficacy/effectiveness through a scientifically validated study?

This can be determined by evaluating whether a research study about the app has been published in a peer-reviewed journal. The research should be specific to the mental health component of the app.

- Yes, there is strong research support for the app (i.e., at least one published paper in a peer-reviewed journal that uses a randomized trial that shows efficacy or effectiveness)
- Yes, there is some research support for the app (i.e., at least one published paper in a peer-reviewed journal that uses single-case design, quasi-experimental methods demonstrating efficacy)
- No, there is no research support for the app

9. Is the app based on or does it use an evidence-based strategy, such as Cognitive Behavioral Therapy (CBT), Dialectical Behavior Therapy (DBT), or evidence-based guidelines?

Evidence-based medicine aims to assess the strength of proof behind medical interventions in terms of risks and benefits and, therefore, can be used to inform clinical decisionmaking on both an individual and a population basis.

- Yes, the app reports using an evidence-based strategy/guidelines to achieve its goals
- No, the app does not report the use of evidence-based strategy/guidelines to achieve its goals (or there are no goals described)
- Unable to assess

D. Linkage to Care

This category evaluates the linkages to healthcare providers, caregivers. This information can be found within the app/app website.

10. Does the app facilitate remote monitoring of the patient or send alerts to a clinician/clinical care team or caregiver?

This question seeks to assess whether a healthcare provider or caregiver can monitor the user's health data in real time AND/OR receive alerts triggered by certain pre-specified high-risk events. Functionality to export data into a file or an email does not count as facilitation of remote monitoring.

- Yes
- No
- Unable to assess

10.a. If YES, please specify *who* can monitor the health of the patient, either through alerts or by other means?

- Healthcare provider only
- Caregiver only
- Both healthcare provider and caregiver
- Other
- Not applicable

E. Access to Crisis Services

This section evaluates whether the app provides access to emergency sources of information.

This information can be found within app/app website. There should be a caregiver or healthcare provider interface within the app. Exporting data into a file or an email would not be considered facilitation of remote monitoring.

11. Does the app provide users with information about resources that can be reached in case of an emergency? This information should be available on the mobile app. Reviewers should not look for this information on the website.

- Yes (e.g., a crisis hotline, 911, nearest emergency room services.)
- No, the app provides no information on emergency services
- Unable to assess

12. Evidence and Linkage to Care: If the app can be used by individuals that require substituted consent OR by minors, then is consent sought from either a caregiver/parent/legal guardian?

This should be assessed at the time of creating a login for the app.

- Yes
- No
- Unable to assess
- Not applicable

Risk Assessment

What is the Risk Category for the App?

A lower number signifies a lower risk level.

Risk Level 1: Apps rated as Risk Level 1 do not target a vulnerable population AND do not provide standalone treatment. These are likely to be apps that provide functions such as mindfulness, meditation, and wellness resources. No specific safety checks are needed for these apps.

- Risk Level is 3:** Response to EITHER Question 5 or Question 6 is **Yes** AND response to Question 7 is **Yes**, go to *Safety Assessment Step 3*
- Risk Level is 2:** Response to EITHER Question 5 or Question 6 is **Yes** OR response to Question 7 is **Yes**, go to *Safety Assessment Step 3*
- Risk Level is 1:** None of the above, go to *Safety Assessment Step 3*

Safety Assessment

Step 1: Preliminary Safety Assessment for Risk Levels 2 & 3

Risk Level 2	<input type="radio"/> [If responses to Questions 9-11 are Yes AND response to Question 10 is Yes or No , supporting evidence-base available]	PRELIM SECURITY CHECK PASSED , go to <i>Safety Assessment Step 2</i>
	Else	PRELIM SECURITY CHECK FAILED , go to <i>Safety Assessment Step 2</i>
Risk Level 3	<input type="radio"/> [If responses to Questions 8-11 are Yes , supporting evidence-base available]	PRELIM SECURITY CHECK PASSED , go to <i>Safety Assessment Step 2</i>
	Else	PRELIM SECURITY CHECK FAILED , go to <i>Safety Assessment Step 2</i>

Step 2: Safety Assessment for Vulnerable Population

If response to Question 5 and/or Question 6 is Yes or Unable to Assess,

- If response to Question 12 is **Yes**, then **SAFE for vulnerable populations**, go to *Safety Assessment Step 3*
- If response to Question 12 is **No or Unable to Assess** then **NOT SAFE for vulnerable populations**, go to *Safety Assessment Step 3*

else go to *Step 3 under Safety Assessment*

STEP 3: Safety Assessment

RISK LEVEL	PRELIM SAFETY CHECK	VULNERABLE POPULATION	APP SAFETY
Risk Level 1	N/A	N/A or SAFE for Vulnerable Populations	Passed Safety Check
	N/A	Not Safe for Vulnerable Populations	Failed Safety Check
Risk Level 2	PASS	N/A or SAFE for Vulnerable Populations	Passed Safety Check
		Not Safe for Vulnerable Populations	Failed Safety Check
	FAIL	--	Failed Safety Check
Risk Level 3	PASS	N/A or SAFE for Vulnerable Populations	Passed Safety Check
		Not Safe for Vulnerable Populations	Failed Safety Check
	FAIL	--	Failed Safety Check

Section 2: Function

This section is focused on descriptive aspects, such as accessibility, usability, privacy/security, AI, cultural competence, and informed consent.

A. Accessibility Features

Accessibility features are meant to make the app easier to navigate, especially for those with disabilities. Common accessibility features include text-to-speech, closed-captioning, and keyboard shortcuts.

1. Which of the phone's accessibility features work within the app?

Reviewers will have to activate accessibility features on their phone under Settings in order to test this.

Select all that apply:

- Text adjustment feature
- Colorblind color scheme feature
- Text-to-speech feature
- None of the phone's accessibility features work in the app
- Unable to assess, please provide a reason: _____

2. Are there additional accessibility features that are provided by the app?

These are accessibility features that are provided within the app but cannot be set as an option under Settings.

Select all that apply:

- Adaptation of audio/video content with transcriptions or captions
- Tapping and other gestures are configurable
- Contrast text coloring in the content
- Screen reader
- No additional accessibility features provided by the app
- Other features not provided here (please specify): _____
- Unable to assess, please provide a reason: _____

B. App Information

General info about the app and user ratings.

The information requested in these questions can be obtained by going to the appropriate app store using your browser.

3. Does the app work on Apple(iOS)?

- Yes
- No

4. What are the number of reviews on the iOS App Store?

The iOS App Store can be accessed through a browser on your phone, tablet, or computer.

5. What is the app rating (number of stars) on Apple Store?

This provides users' perspectives of the app.

6. Does the app work on Android?

- Yes
- No

7. What are the number of reviews on Google Play Store.

8. What is the app rating (number of stars) on Google Play Store?

The Google Play Store can be accessed through a browser on your phone, tablet, or computer.

C. Costs

App costs could be upfront, in the form of a monthly/annual subscription, or through freemium services that require in-app purchasing. Some apps require payment from the user, while others are reimbursed by the healthcare provider or insurance.

The information requested in these questions may be found in the app, in the App Store, and on the app site.

9. What is the business model for the app? More than one option may be applicable here. An app that provides a free trial for a limited duration is not considered to be free.

If the app provides a payment gateway to pay a therapist or healthcare provider, those costs should not be included.

- Free (no upfront fee for the app; additional packages/services may be offered for a fee)
- Upfront fee (a onetime cost for accessing the app; additional packages/services may be offered for additional fees)
- In-app purchases (additional packages/services available for a fee in addition to or in lieu of an upfront cost, e.g., concierge services)
- Subscription (payment for services on a monthly/quarterly/annual basis)
- Reimbursed by healthcare providers/insurers/employers
- Other, please specify: _____

10. Does the app provide a free or freemium model? An app that provides a free trial for a limited duration is not considered to be free.

- Free
- Freemium
- No free or freemium version

11. What is the estimated annual cost of the app for the paid version?

If the app provides a free or freemium model, then select "Not applicable" in the list below.

- Under \$50
- \$50-250
- Over \$250
- Not applicable

12. If the app includes paid service(s), does it provide CPT code(s) for insurance reimbursement?

Current Procedural Terminology (CPT) are unique 5-digit codes for medical services or procedures assigned by the American Medical Association which are used throughout the U.S. medical system. Insurance companies decide on which CPT codes can be used for insurance reimbursement.

- Yes
- No or the app has no paid services
- Unable to Assess, please provide a reason: _____

D. Organizational Credibility

Organizational credibility is meant to determine whether the app comes from a trusted source.

13. Who is the developer of the app?

This information can be found on the app website or in the Terms of Service on the app. More than one response may apply. For example, two or more institutions may be involved in the development of the app; or, for example, if an app was developed by Johns Hopkins University, then "Non-profit institution" and "Academic institution" would apply.

- Government
- For-profit company
- Non-profit institution
- Healthcare institution
- Academic institution
- Insurance company
- Independent developer/s
- Unable to assess
- Other, please specify: _____

14. Does the app have any consumer bureau complaints or lawsuits pending?

This information can be found here: Better Business Bureau: <https://www.bbb.org/us/ca/san-francisco/profile/mobile-apps/calmcom-inc-1116-877519/complaints> ; Google Search: <Consumer Complaints> <app name> <app/mobile app>; Google Search: <law suits> <app> <app/mobile app>

- Yes
- No
- Unable to assess

E. Evidence & Clinical Foundation

The information requested in these questions can be found on the app website.

15. Does the app appear to do what it claims to do?

Look at the website to see the claims made by the app and compare them to what you see in the app.

- Yes, the app provides the functionality it claims on its website
- The app provides some of the functionality it claims on its website
- No, the app does not provide the functionality it claims on its website
- Unable to assess

F. Privacy & Security

Security is about the safeguarding of data and unauthorized access of the app. Privacy is about the safeguarding of the user's identity.

The information requested in these questions should be available in the app in the Terms & Conditions or Privacy & Security agreement. Do not look at information provided on the website.

16. Does the app claim it meets HIPAA [or analogous national standard for protected health information (PHI)]?

- Yes
- No

17. Does the app claim it meets COPPA [or analogous national standard for protected health information (PHI) for minors younger than 13 years of age]?

The Children's Online Privacy Protection Act (COPPA) applies to the online collection of personal information by persons or entities under U.S. jurisdiction about children younger than 13 years of age.

- Yes
- No

18. Does the app report sharing or selling of data for research or commercial purposes?

Information used internally within the app itself to make better recommendations to the user does not count as sharing or selling. Examples of apps that may be sharing or selling data include those that provide evidence based drug and dosage information, clinical decision support systems where you share conditions, symptoms and medication information, medication adherence apps, etc.

- Yes
- No
- Unable to assess

19. If your response to question 18 is “YES”, please answer the following question:

If the app reports sharing or selling for research or commercial purposes, are the data de-identified?

This question is targeted at understanding if the data is de-identified or not. Information used internally, within the app itself, to make better recommendations to the user does not count as sharing or selling.

- Yes, the app reports that data is de-identified
- No, the app does not report data is de-identified
- Unable to assess

20. If the app has the capability to read/write to an electronic health record management system (EHRs) or other healthcare systems, does it use industry standards for secure interoperability (e.g., FHIR, SMART, OAuth 2.0, TLS 1.2)?

This information may need to be deduced. If the app has existing integrations with known EHRs such as EPIC, it can be concluded that it follows industry standards. This information may also need to be identified through a review of the app developer webpage.

- Yes, the app uses industry standards for interoperability
- No, the app does not use industry standards for interoperability
- Not applicable, because it does not read/write to EHRs
- Unable to assess

G. Informed Consent

Informed consent is permission granted by the user after information about the potential use or disclosure of the information collected by the app is provided. Most apps tend to have a disclosure list that is long and hard to understand. There are best practices for ensuring that users understand exactly what they are agreeing to before they “Agree” to the app’s privacy and security practices. The following questions evaluate whether the app follows these best practices.

The information requested in these questions can be found within the app.

21. Please assess the level of informed consent enabled by the app.

Providing consent is often one of the first things users are asked to do after downloading an app. Traditional written consent forms (e.g., Terms of Service, Privacy and Security agreement) are often too long, difficult to understand, and overly focused on legal concerns. It is important that apps enable informed consent for prospective users in a manner that is easy to understand.

- No informed consent
- Does not simplify informed consent
- Average: Focuses on the essential by a) providing a narrative focused on the most salient information, b) limiting concepts to one per screen, and c) following national plain language and health literacy guidance
- Good: Focuses on the essential and organizes content deliberately by a) prioritizing key words/concepts presented on each screen, b) providing information tiers for conceptual elaboration, and c) enabling participants to navigate to their desired level of detail
- Excellent: Focuses on the essential, organizes content deliberately, and encourages engagement through interactive elements

22. What is the format of the data privacy and security consent process followed in the app?

*An opt-in consent process requires **explicit consent from the user** before the collection and processing of their personal data. It refers to a **positive action** taken by the user indicating that they agree to the use of their personal data. An opt-out consent process **does not require the user’s consent** prior to the collection and processing of their personal data. It refers to the process by which users **withdraw their consent** to the use of their personal data.*

- Opt-in for data to be shared, the default choice is opt-out
- Opt-out of data sharing, the default choice is opt-in
- No choice to opt-out of data sharing, without explicit consent you can’t use the app
- Unable to assess

H. Cultural Competence

Cultural competence is defined as the ability to understand, appreciate, and account for different cultures or belief systems based on race, ethnicity, income strata, religious beliefs, etc. It is important to assess if the app captures these differences and whether it provides personalized care that takes these cultural differences into account.

The information requested in these questions can be found within the app or on the app website.

23. Does the app report developing and testing the app for specific cultural group/s?

According to the U.S. Department of Health and Human Services (HHS), "cultural competence is the integration and transformation of knowledge about individuals and groups of people into specific standards, policies, practices, and attitudes used in appropriate cultural settings to increase the quality of services; thereby producing better outcomes." For this assessment, groups with lived experiences such as pregnant teens and survivors of gender-based violence would be considered to be a specific cultural group.

- Yes
- No
- Unable to assess
- Other, please specify: _____

24. If the app reports developing and testing the app for specific cultural groups, is there published documentation (e.g., website, papers) of the process taken to incorporate information that is specific to specified cultures?

Select all that apply:

- Website
- Documentation (on the app or website)
- Published scientific papers
- Unable to assess
- Other, please specify: _____
- Not applicable

If the app integrates culturally specific groups, please name the groups.

25. Is gender-inclusive language employed? When asking the user's gender, is there an option to self-describe, an option to decline to answer, use of scientifically correct terms for gender (e.g., man, woman, nonbinary)?

This information can be found while assessing the app and is generally encountered when a user creates an account on the app.

Select all that apply:

- An option for use of personal pronouns (e.g., they/them, he/his, she/her)
- An option to specify another gender
- An option to decline to answer
- Use of scientifically correct terms for gender, such as man, woman, non-binary
- Unable to assess
- Not applicable

26. If the app was tested in a study, what was the percentage of non-white participants?

For clinical trials, this information may be found on the clinicaltrials.gov page of the trial under Study Results (include "Other" as non-white). If the study was a usability or other type of study, this information may be found on the developer website.

- Less than 30%
- Between 30 – 50%
- More than 50%
- No disaggregate data available
- No information available

I. Usability

Usability can be described as the capacity of a system to provide a condition for its users to perform the tasks safely, effectively, and efficiently. It is important that the experience be engaging and pleasing, otherwise users are likely to stop using the app.

The information requested in these questions can be found within the app.

27. Does the app work offline?

Choose Yes if parts of the app or the entire app work offline.

- Yes
- No
- Unable to assess

28. Performance: How accurately/fast do the app features (functions) and components (buttons/menus) work?

- App is broken; no/insufficient/inaccurate response (e.g., crashes/bugs/broken features)
- Some functions work, but app lags or contains major technical problems
- App works overall; some technical problems need fixing or app is slow at times
- Mostly functional with minor/negligible problems
- Perfect/timely response; no technical bugs found or contains a 'loading time left' indicator, if relevant

29. What languages are supported by the application?

This information can be found in the App Store.

Select all that apply:

- English
- Spanish
- French
- German
- Other, please specify: _____

30. Customization: Does the app allow customization of settings and preferences by the user (e.g., sound, content, notifications)?

- Does not allow any customization or requires setting to be input every time
- Allows little customization and that limits the app's functions
- Basic customization to function adequately
- Allows numerous options for customization
- Allows complete tailoring to the user's characteristics/preferences, remembers all settings

31. Target group: Is the app content (visuals, language, design) appropriate for the target audience?

- Completely inappropriate, unclear, or confusing
- Mostly inappropriate, unclear, or confusing
- Satisfactory, but not specifically designed for the target audience, may be inappropriate/unclear/confusing at times
- Designed for the target audience, with minor issues
- Designed specifically for the target audience; no issues found

32. Layout: Is the arrangement and size of buttons, icons, menus, and content on the screen appropriate?

- Very bad design: cluttered, some options impossible to select/locate/see/read
- Bad design: random, unclear, some options difficult to select/locate/see/read
- Satisfactory: few problems with selecting/locating/seeing/reading items
- Mostly clear: able to select/locate/see/read items
- Professional: simple, clear, orderly, logically organized

33. Graphics: How high is the quality/resolution of graphics used for buttons, icons, menus, and content?

- Graphics appear amateur, very poor visual design: disproportionate, stylistically inconsistent
- Low quality/low resolution graphics, low quality visual design: disproportionate
- Satisfactory quality graphics and visual design: generally consistent in style
- High quality/resolution graphics and visual design: mostly proportionate, consistent in style
- Very high quality/resolution graphics and visual design: proportionate, consistent in style throughout

34. Visual appeal: How good does the app look?

- Ugly/unpleasant to look at: poorly designed, clashing, mismatched colors
- Bad: poorly designed, bad use of color, visually boring
- Satisfactory: average, neither pleasant nor unpleasant
- Pleasant: seamless graphics, consistent and professionally designed
- Beautiful: very attractive, memorable, stands out; use of color enhances app features/menus

35. Does the app have advertising?

- Yes
- No

36. If app has advertising, is the advertising intrusive and distracting?

- Advertising is neither intrusive nor distracting
- The advertising is both intrusive and distracting
- Not applicable

37. Ease of use: How easy is it to learn how to use the app; how clear are the menu labels, icons, and instructions?

- No/limited instructions, menu labels and icons are confusing, complicated
- Takes a lot of time or effort
- Takes some time or effort
- Easy to learn (or has clear instructions)
- Able to use app immediately, intuitive and simple (no instructions needed)

38. Navigation: Does moving between screens make sense? Does the app have all necessary links between screens?

- No logical connection between screens at all/navigation is difficult
- Understandable after a lot of time/effort
- Understandable after some time/effort
- Easy to understand/navigate
- Perfectly logical, easy, clear and intuitive screen flow throughout, and/or has shortcuts

39. Gestural design: Do taps/swipes/pinches/scrolls make sense? Are they consistent across all components/screens?

- Completely inconsistent/confusing
- Often inconsistent/confusing
- Satisfactory with some inconsistencies/confusing elements
- Mostly consistent/intuitive with negligible problems
- Perfectly consistent and intuitive

40. Content: Is the app copy well written and relevant to the goal and/or topic of the app?

- There is no information within the app
- Irrelevant/inappropriate/incoherent/incorrect
- Barely relevant/appropriate/coherent, may be incorrect
- Satisfactory with respect to relevance/appropriateness/coherence and appears to be correct
- Relevant/appropriate/coherent/correct
- Highly relevant/appropriate/coherent/correct
- Not applicable

41. Is there any evidence to show the average duration of use of the app by users?

This information may be available on the app website or in a paper published in a scientific journal. Do not look at user reviews on the app website, in the App Store, or in news articles.

- On average, users have engaged with the app for 30 days or less
- On average, users have engaged with the app for more than a month but less than 6 months
- On average, users have engaged with the app for more than 180 days
- No evidence available

J. Remote Monitoring

Remote patient monitoring enables monitoring of patients outside of conventional clinical settings through real-time access to patient health data. The patient's provider may have a schedule to monitor patient data or may receive alerts about their patient's health.

The information requested in these questions can be found within the app/app website.

42. If remote monitoring is feasible, how does the provider access the data?

- Provider has access to patient information through the app
- Provider has access to the patient information through the EMR
- Provider has no access to patient information
- Not applicable
- Unable to assess
- Other, please specify: _____

43. Does the app provide alerts to the provider to notify them of a clinical event that may require action on their part?

Two-way messaging capability with a healthcare provider or concierge may be considered to be an alerting mechanism.

- Yes
- No
- Unable to assess

44 . Can the app share data with wearables like Apple Watch and Fitbit?

This option will either pop up on the app during sign-up or be mentioned on the app developer's website.

- Yes
- No
- Unable to assess

K. Access to Crisis Response Services

This section evaluates whether the app provides access to emergency sources of information.

The information requested in this question can be found within the app or on the app website.

45. Does the app connect the user automatically to resources in case of a crisis situation or emergency?

- Yes
- No
- Unable to assess

L. Artificial Intelligence (AI)

Apps often use artificial intelligence (AI) for a variety of tasks including risk-prediction, differential diagnoses, and personalization of health content. The following questions are important because we want to gauge the potential of the app to cause harm and also to determine whether the app's algorithms are being updated based on user input.

The information requested in these questions can be found in the app, on the app website, or from evidence found through a literature search.

46. Does the app use AI? If yes, please respond to questions 47 and 48.

- Yes
- No

47. How is AI being used in the app?

For example, AI may be used in the app to automate simple tasks, such as scheduling, or it may be used for more complex functionality, such as to personalize chatbots.

48. Is there any evidence to suggest that the app uses data from user interactions to improve precision on AI models?

For example, after an interaction, the app may ask the user for feedback on the automated response from the app to see if it was appropriate. If the response was not correct, the app may ask for further feedback from the user so the algorithms can be refined.

- Yes, there is evidence to suggest that the AI models are updated based on feedback
- No, there is no evidence to suggest that the AI models are updated based on feedback
- Unable to assess

Section 3: Mental Health App Features

This section attempts to capture the features specific to mental health present in the app. This is not meant to be an exhaustive list. Only some of these features may apply to the app being reviewed.

The information requested in these questions can be found in the app, on the app website or in published literature.

1. List mental health symptom(s) and/or condition(s) addressed by the app.

2. Please answer the following questions about the app features.

Question	Yes	No	Unable To Assess
a. Does the app facilitate text messaging interactions with healthcare therapists, coaches, or other providers via the app?			
b. Does the app facilitate audio chat interactions with healthcare therapists, coaches, or other providers via the app?			
c. Does the app facilitate video chat interactions with healthcare therapists, coaches, or other providers via the app?			
d. Does the app facilitate teletherapy services via the app?			
e. Does the app facilitate group therapy services via the app?			
f. Does the app provide live support to a coach or counselor via the app?			
g. Does the app provide concierge mental health services via the app? Concierge services are personalized services patients can choose based on what they think works best for them. Some apps may provide acute concierge services to include additional help during times of crisis or high stress.			

3. Does the app provide a direct connection to 988 or other hotlines?

- Yes, and the connection works
- Yes, but the connection does not work
- No
- Unable to assess

4. Are the following functionalities supported by the app?

The following questions ask you to rate the functionality as either “comprehensive” or “not comprehensive.”

Comprehensive: Broad range of functionality that meets or exceeds the needs of the user and allows for flexibility

Not comprehensive: Rudimentary functionality that may not meet the needs of the user or provides limited choices and lacks flexibility

	Comprehensive	Not Comprehensive	Not Applicable
<p>Mindfulness: Mindfulness is a therapeutic technique that includes elements of relaxation, breathing, and body exercise. It also includes techniques, such as meditation, guided positive imagery, grounding exercises, or progressive muscle relaxation.</p>			
<p>Journaling: Journaling can help users manage anxiety, reduce stress, and cope with mental health challenges. It can provide users with a way of identifying negative thoughts and behaviors and highlight positive aspects in their lives.</p>			
<p>Psychoeducation: Psychoeducation may be defined as the education of a person with a psychiatric disorder regarding the symptoms, treatments, and prognosis of that illness. Psychoeducation may also be targeted to the caregiver, family member, or loved one of the patient. Psychoeducation should be a brief personal intervention by a healthcare provider upon first diagnosis; however, sometimes, only written materials or online resources are provided to the patient. In the worst case scenario, patients receive no Psychoeducation from their provider. (Source: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD010823.pub2/full)</p>			
<p>Skill Building: Skill building includes techniques for recognition of signs and symptoms of a problem, self-advocacy, stress management, emotional regulation, relapse prevention, promotion of sleep hygiene, etc. The app may include tips and advice on dealing with negative emotions and with behavior change. Skill building may require repetition, positive reinforcement, modeling, and practice.</p>			
<p>Screening: Historically, mental health care has relied on structured patient interviews and self-reported questionnaires for diagnosis. These can be used for self-evaluation, reporting, or to decide whether a patient should engage with a mental health professional. Please indicate if industry-validated screening questionnaires (e.g., GATT, PHQ-4, PHQ-9) are provided by the app. The standardized screening questionnaires used will generally be mentioned on the app developer's website.</p>			
<p>Safety Planning: Safety planning is designed to help individuals respond to escalation of suicidal thoughts and feelings, giving them a tailored list of coping behaviors and social support to use until the risk passes.</p>			

	Comprehensive	Not Comprehensive	Not Applicable
Sleep Hygiene: This includes features for promoting sleep hygiene, such as sleep diaries, lifestyle tracking, alarms, data synchronization with a wearable device (e.g., FitBit, Apple Watch), etc.			
Automated Chatbots: Chatbots include conversational agents that can provide virtual therapy and/or social support.			
Family/Caregiver Support: Caregivers take care of the day-to-day needs of patients. Apps can help caregivers monitor the mental state of the patient, identify the early signs of illness, track relapse and deterioration, and help the patient access services. The family/caregiver may also supervise treatment and provide emotional support to the patient.			
Social and Peer Group Interaction: Individuals with mental health conditions may find it hard to engage socially and may have a small social network/peer group. Apps may provide peer group interactions and/or community support discussion boards for such individuals.			
Gamification: Gamification is the application of typical elements of game playing (e.g., point scoring, competition with others, rules of play) to other areas of activity. Gamification is used to encourage and reward positive changes.			
Personalization: Personalized treatment entails the selection of a therapy or treatment protocol based on a patient's profile, which may increase the likelihood of a successful outcome. Personalization can be used to tailor interactions initially, when the user starts to use the app, and also on a more regular basis. It is usually based on user information collected by the app, such as what activities produce positive emotions (e.g., physical exercise, talking to a friend).			
Self-Monitoring: Self-monitoring apps allow users to engage in symptom monitoring, mood tracking, and progress tracking through a treatment program, such as one for substance abuse, depression and anxiety, etc.			
Medication Adherence: Does the app support patients in their use of medication(s)? This could be through alerts, notifications, or other means.			
Medication Delivery: Does the app support physical delivery of medications to patient(s)?			

5. Please add any features supported by the app that are not listed above.

Post Administrative Questionnaire

This section contains some details about the app and evidence being reviewed.

1. Include all links to references for evidence (citations on PubMed, systematic reviews, websites, etc.) here.

2. Based on your review of the app, do you think it would have been useful to have some training or a tutorial about how to use it?

This is your subjective assessment based on your usage of the app.

- Yes, there was training available, and I needed it to use the app
- Yes, there was training available, but I didn't need it to use the app
- No, there was no training available, but I needed it to use the app
- No, there was no training available, and I didn't need it to use the app
- Don't know

3. Does the app have any help-related documentation available in the app itself (e.g., tooltips, general help)?

Base your answer on what was available on the app itself.

In-app help usually comes in three forms: 1) a separate page or pages of help within the app, with instructions; 2) popups that provide contextual help, displaying instructions relevant to the specific task that the user is attempting; 3) descriptions of app features of interest to the user. If at least two of these features is present, the help-related documentation would be considered to be comprehensive.

- Yes, there was comprehensive help available on the app
- Yes, there was some help available on the app
- The app help button took me to the website
- No help was available on the app

4. Does the app have a dedicated website that provides information about the app?

This may be part of a company website or a standalone website focused solely on the app.

- Yes
- No
- Unable to assess

5. Does the app have any help-related documentation available on its website?

- Yes, there was comprehensive help available
- Yes, there was some help available
- No, help was not available on the app

6. Please provide a subjective evaluation of the app in your own words.

- Do you think the assessment conducted here matches your subjective assessment? If not, please provide additional details.
 - Do you think the risk posed by the app is captured accurately here?
 - Do you think the app is technically sound?
 - Do you think the therapeutic content provided by the app is rudimentary or substantive?
 - Can this app help those that may be experiencing mental or behavioral health challenges?
-

7. Do you think the app could cause harm or have a negative impact on the user? Please elaborate.

8. Do you think the subjective risk assessment matches the risk assessment calculated in Section 1? Please elaborate.

9. Any additional comments?

10. Please note the time taken to complete your assessment using this framework.

Glossary

Term	Explanation
Artificial Intelligence (AI)	The ability of a computer to do tasks that are usually done by humans.
Caregiver	A family member or helper who regularly looks after a child or a sick, elderly, or disabled person.
Cultural Competence	The ability to understand, appreciate, and account for different cultures or belief systems based on race, ethnicity, income strata, religious beliefs, etc.
Functional Impairment	Limitations owing to illness(es) that cause an inability to carry out certain functions in one's daily life.
Gamification	Application of typical elements of game playing (e.g., point scoring, competition with others, badges and awards) to other areas of activity to encourage engagement with a product or service.
Informed Consent	Permission granted in the knowledge of the possible consequences, typically that which is given by a patient to a doctor for treatment with full knowledge of the possible risks and benefits.
Vulnerable Population	For the purposes of this report, either minors (younger than 18 years of age) or adults with a level of functional impairment that causes them to need support from a caregiver.

Appendix G. Framework Assessment Examples

We provide examples of using the FASTER Framework to assess three apps:

Assessment Example, App A – Poor quality app

Assessment Example, App B – Medium quality app

Assessment Example, App C – High quality app

Assessment Example, App A

Administrative Questionnaire

This section contains introductory questions about the app to be reviewed. The questions are mostly factual, rather than evaluative. Some of these questions may make it into the final Assessment Framework, based on the information we gather from the pilot.

Q 1. Reviewer Name

Q 2. Date of Evaluation

Q 3. App Name

Q 4. App/Company Website

This should only be a specific website developed by the app developer and not their pages on social media sites, such as Instagram or Facebook.

Q 5. Country of Origin

This information can be found in the Terms & Conditions on the website or in the app.

Q 6. App Version Number

This information can be found on the app and from the App Store.

Q 7. Platform Used for Assessment by the Reviewer

Q 8. OS Version Number on Evaluating Device

For Android go to: Settings > About Phone > Software Information;
For iOS go to: Settings > General > Software Version.

n/ a

Q 9. Describe the goal of the app in your own words. Do not use any of the marketing phrases/lingo used by the app developer.

This should be based on the app description on the website.

to provide information on bipolar disorder

Q 10. Has the app been approved by any regulating authority, such as FDA?

This information can be found on the app developer website.

no

Q 11. Does the app have multiple revenue models, such as freemium, in-app purchases, etc.?

- Yes
- No

Q 12. If the app has multiple versions (e.g., free, freemium, paid), which version did you evaluate? If the app had only one revenue model, please choose “not applicable” here.

- Free
- Paid version
- Paid by insurance agency, employer, or healthcare provider
- Not applicable
- Unable to assess

Q 13. Does the app require prior authorization from a healthcare institution/insurance provider/college/employer or other institution for access?

- Yes
- No
- Unable to assess

Q 14. If you were unable to assess the app, please specify the reason.

n/a

Note: Questions about privacy, security, and informed consent are covered in both Section 1 and Section 2 of the framework. Please make a note of all the disclaimers, warnings, and privacy and security statements and agreements when you first log into the app, so you are able to answer the questions that come later in the framework.

Section 1: Risks and Mitigation Strategies

Section 1 aims to evaluate the potential risk posed by the app, and evidence of efficacy, and safety features. The burden of evidence to prove efficacy and safety is dependent on the goals of the app, the target audience, and the severity of the mental health condition.

The categories of questions in this section are App Integrity, Risk Assessment, Evidence, Linkage to Care, and Access to Crisis Services. If apps do not meet the threshold for safety and credibility, that is flagged in the assessment.

A. App Integrity

The information requested in these questions can be found on the App Store, the app's website, or via Google Search.

1. Was the version of the app you are reviewing updated in the last 6 months?

- Yes
- No
- Unable to assess

2. Does the app provide a privacy and security agreement to be reviewed and agreed to by the user?

The agreement should be accessible through a link on the app or on the mobile app. Reviewers should not look for this information in other locations, such as the developer's website.

- Yes, there is a privacy and security agreement to be agreed to by the user
- No, there is not a privacy and security agreement
- Unable to assess or unable to access

3. Does the app provide warnings and disclaimers (e.g., the limitations of the app related to medical liability)?

Reviewers should look for this information on the mobile app ONLY.

- Yes, warnings and/or disclaimers are provided
- No, warnings and/or disclaimers are not provided
- Unable to assess
- Not applicable

4. Has the app been endorsed by, or is it being used by, a government agency or trusted mental health professional association?

Endorsement and/or usage helps us establish the credibility of the app.

Endorsement means to give one's approval to, especially officially by an organization. An app merely being reviewed by the American Psychological Association (APA), or a similar organization, does not mean it has been approved by them. Endorsement should be for the app and not the company developing the app or the domain (e.g., teletherapy). Self-reported endorsements are acceptable.

Usage by an institution signifies that an app is made available to all its members. For example, Veterans Affairs uses iBlueButton to deliver critical health care information to veterans.

Examples of government agencies include federal, state and local agencies such as CMS, Veterans Affairs, Virginia Department of Health. Examples of trusted mental health professional associations include Substance Abuse and Mental Health Services Administration (SAMHSA), American Psychological Association (APA), and National Alliance on Mental Illness (NAMI).

- Yes, the app has been endorsed by one or more mental health associations, government agencies, or non-government bodies (examples may include but are not limited to government agencies, such as Centers for Medicare & Medicaid Services (CMS) or Veterans Affairs (VA); non-government bodies, such as APA; and insurance or healthcare institutions)
- No, the app has not been endorsed
- Unable to assess, please specify reason: _____

App Integrity Assessment

Question	Response	App Integrity Level
Does the app satisfy the Integrity Assessment? Please pick only 1 response.	<input type="radio"/> Responses to questions 1-3 are all Yes OR response to question 4 is Yes	App Integrity is "High"
	<input checked="" type="radio"/> Response to any of the questions 1-3 is No AND response to 4 is No	App Integrity is "Low"

B. Risk Assessment

This category aims to determine the level of risk posed by the app. For example, if the app is used by someone with a clinically diagnosed condition, such as Schizophrenia, then the risk posed by the app may be higher than if the person has mild anxiety. Similarly, if the app targets children and adolescents, then the potential risk associated with the app may be higher than if it targeted adults with no impairment or mental health diagnosis.

*For the purposes of this report, either minors (younger than 18 years of age) or adults with a moderate or severe level of functional impairment that causes them to need support from a caregiver are considered to be part of a **vulnerable population**.*

The information requested in these questions can be found within the app/app website.

5. Is the app intended for use by adults who may have an illness or disability that impacts their decisionmaking ability? Severe autism or severe dementia are examples of illnesses or disabilities that may affect decisionmaking ability.

For the purposes of this assessment, the population that may have an illness or disability that impacts their decision making ability is included in the "vulnerable population" group. Adults can usually make medical decisions, such as consenting to treatment, on their own. In some cases, adults need someone else to make medical decisions for them if their decisionmaking ability is affected by an illness or disability. This is sometimes referred to as "substituted consent." A higher level of evidence/protection is needed for the use of specific apps under these circumstances.

- Yes
- No
- Unable to assess

6. Does the app claim that it is intended for use by minors (i.e., people younger than 18 years of age)?

For the purposes of this assessment, this population is included in the "vulnerable population" group. The app is "intended for use" by a population if it is clearly specified on the app or website that this population is its target user group. Do not look at the App Store Age Ratings or Terms and Conditions for this information.

- Yes, for use by minors
- Yes, for use by both minors and adults
- No, not for use by minors
- Unable to assess

7. Does the app claim to provide standalone treatment for any mental health condition?

A standalone treatment might be some form of psychotherapy or other any other medical care provided by the app independently, rather than in conjunction with a licensed healthcare practitioner.

Functionality, such as mindfulness training, skills training, symptom or mood tracking, and social support, would not be considered standalone treatment.

- Yes, the app provides standalone treatment

- No, the app does not provide standalone treatment
- Unable to assess

C. Evidence

It is important that the app has a solid clinical foundation. If the app targets a vulnerable population, the burden of evidence to prove efficacy and safety is greater.

The information requested in these questions can be found on the app website (look at customer testimonials or resources/blog for anecdotal information) and on PubMed (search PubMed as outlined in the training document). Make sure to save all the references in the app review document.

8. Has the app been evaluated for efficacy/effectiveness through a scientifically validated study?

This can be determined by evaluating whether a research study about the app has been published in a peer-reviewed journal. The research should be specific to the mental health component of the app.

- Yes, there is strong research support for the app (i.e., at least one published paper in a peer-reviewed journal that uses a randomized trial that shows efficacy or effectiveness)
- Yes, there is some research support for the app (i.e., at least one published paper in a peer-reviewed journal that uses single-case design, quasi-experimental methods demonstrating efficacy)
- No, there is no research support for the app

9. Is the app based on or does it use an evidence-based strategy, such as Cognitive Behavioral Therapy (CBT), Dialectical Behavior Therapy (DBT), or evidence-based guidelines?

Evidence-based medicine aims to assess the strength of proof behind medical interventions in terms of risks and benefits and, therefore, can be used to inform clinical decisionmaking on both an individual and a population basis.

- Yes, the app reports using an evidence-based strategy/guidelines to achieve its goals
- No, the app does not report the use of evidence-based strategy/guidelines to achieve its goals (or there are no goals described)
- Unable to assess

D. Linkage to Care

This category evaluates the linkages to healthcare providers, caregivers. This information can be found within the app/app website.

10. Does the app facilitate remote monitoring of the patient or send alerts to a clinician/clinical care team or caregiver?

This question seeks to assess whether a healthcare provider or caregiver can monitor the user's health data in real time AND/OR receive alerts triggered by certain pre-specified high-risk events. Functionality to export data into a file or an email does not count as facilitation of remote monitoring.

- Yes
- No
- Unable to assess

10.a. If YES, please specify *who* can monitor the health of the patient, either through alerts or by other means?

- Healthcare provider only
- Caregiver only
- Both healthcare provider and caregiver
- Other
- Not applicable

E. Access to Crisis Services

This section evaluates whether the app provides access to emergency sources of information.

This information can be found within app/app website. There should be a caregiver or healthcare provider interface within the app. Exporting data into a file or an email would not be considered facilitation of remote monitoring.

11. Does the app provide users with information about resources that can be reached in case of an emergency? This information should be available on the mobile app. Reviewers should not look for this information on the website.

- Yes (e.g., a crisis hotline, 911, nearest emergency room services.)
- No, the app provides no information on emergency services
- Unable to assess

12. Evidence and Linkage to Care: If the app can be used by individuals that require substituted consent OR by minors, then is consent sought from either a caregiver/parent/legal guardian?

This should be assessed at the time of creating a login for the app.

- Yes
- No
- Unable to assess
- Not applicable

Risk Assessment

What is the Risk Category for the App?

A lower number signifies a lower risk level.

Risk Level 1: Apps rated as Risk Level 1 do not target a vulnerable population AND do not provide standalone treatment. These are likely to be apps that provide functions such as mindfulness, meditation, and wellness resources. No specific safety checks are needed for these apps.

- Risk Level is 3:** Response to EITHER Question 5 or Question 6 is **Yes** AND response to Question 7 is **Yes**, go to *Safety Assessment Step 3*
- Risk Level is 2:** Response to EITHER Question 5 or Question 6 is **Yes** OR response to Question 7 is **Yes**, go to *Safety Assessment Step 3*
- Risk Level is 1:** None of the above, go to *Safety Assessment Step 3*

Safety Assessment

Step 1: Preliminary Safety Assessment for Risk Levels 2 & 3

Risk Level 2	<input type="radio"/> [If responses to Questions 9-11 are Yes AND response to Question 10 is Yes or No , supporting evidence-base available]	PRELIM SECURITY CHECK PASSED , go to <i>Safety Assessment Step 2</i>
	Else	PRELIM SECURITY CHECK FAILED , go to <i>Safety Assessment Step 2</i>
Risk Level 3	<input type="radio"/> [If responses to Questions 8-11 are Yes , supporting evidence-base available]	PRELIM SECURITY CHECK PASSED , go to <i>Safety Assessment Step 2</i>
	Else	PRELIM SECURITY CHECK FAILED , go to <i>Safety Assessment Step 2</i>

Step 2: Safety Assessment for Vulnerable Population

If response to Question 5 and/or Question 6 is Yes,

- If response to Question 12 is **Yes**, then **SAFE for vulnerable populations**, go to *Safety Assessment Step 3*
- If response to Question 12 is **No**, then **NOT SAFE for vulnerable populations**, go to *Safety Assessment Step 3*

else go to *Step 3 under Safety Assessment*

STEP 3: Safety Assessment

RISK LEVEL	PRELIM SAFETY CHECK	VULNERABLE POPULATION	APP SAFETY
Risk Level 1	N/A	N/A or SAFE for Vulnerable Populations	Passed Safety Check
	N/A	Not Safe for Vulnerable Populations	Failed Safety Check
Risk Level 2	PASS	N/A or SAFE for Vulnerable Populations	Passed Safety Check
		Not Safe for Vulnerable Populations	Failed Safety Check
	FAIL	--	Failed Safety Check
Risk Level 3	PASS	N/A or SAFE for Vulnerable Populations	Passed Safety Check
		Not Safe for Vulnerable Populations	Failed Safety Check
	FAIL	--	Failed Safety Check

Section 2: Function

This section is focused on descriptive aspects, such as accessibility, usability, privacy/security, AI, cultural competence, and informed consent.

A. Accessibility Features

Accessibility features are meant to make the app easier to navigate, especially for those with disabilities. Common accessibility features include text-to-speech, closed-captioning, and keyboard shortcuts.

1. Which of the phone's accessibility features work within the app?

Reviewers will have to activate accessibility features on their phone under Settings in order to test this.

Select all that apply:

- Text adjustment feature
- Colorblind color scheme feature
- Text-to-speech feature
- None of the phone's accessibility features work in the app
- Unable to assess, please provide a reason: _____

2. Are there additional accessibility features that are provided by the app?

These are accessibility features that are provided within the app but cannot be set as an option under Settings.

Select all that apply:

- Adaptation of audio/video content with transcriptions or captions
- Tapping and other gestures are configurable
- Contrast text coloring in the content
- Screen reader
- No additional accessibility features provided by the app
- Other features not provided here (please specify): _____
- Unable to assess, please provide a reason: _____

B. App Information

General info about the app and user ratings.

The information requested in these questions can be obtained by going to the appropriate app store using your browser.

3. Does the app work on Apple(iOS)?

- Yes
- No

4. What are the number of reviews on the iOS App Store?

The iOS App Store can be accessed through a browser on your phone, tablet, or computer.

n/a

5. What is the app rating (number of stars) on Apple Store?

This provides users' perspectives of the app.

n/a

6. Does the app work on Android?

- Yes
- No

7. What are the number of reviews on Google Play Store.

22

8. What is the app rating (number of stars) on Google Play Store?

The Google Play Store can be accessed through a browser on your phone, tablet, or computer.

4.3

C. Costs

App costs could be upfront, in the form of a monthly/annual subscription, or through freemium services that require in-app purchasing. Some apps require payment from the user, while others are reimbursed by the healthcare provider or insurance.

The information requested in these questions may be found in the app, in the App Store, and on the app site.

9. What is the business model for the app? More than one option may be applicable here. An app that provides a free trial for a limited duration is not considered to be free.

If the app provides a payment gateway to pay a therapist or healthcare provider, those costs should not be included.

- Free (no upfront fee for the app; additional packages/services may be offered for a fee)
- Upfront fee (a onetime cost for accessing the app; additional packages/services may be offered for additional fees)
- In-app purchases (additional packages/services available for a fee in addition to or in lieu of an upfront cost, e.g., concierge services)
- Subscription (payment for services on a monthly/quarterly/annual basis)
- Reimbursed by healthcare providers/insurers/employers
- Other, please specify: _____

10. Does the app provide a free or freemium model? An app that provides a free trial for a limited duration is not considered to be free.

- Free
- Freemium
- No free or freemium version

11. What is the estimated annual cost of the app for the paid version?

If the app provides a free or freemium model, then select “Not applicable” in the list below.

- Under \$50
- \$50-250
- Over \$250
- Not applicable

12. If the app includes paid service(s), does it provide CPT code(s) for insurance reimbursement?

Current Procedural Terminology (CPT) are unique 5-digit codes for medical services or procedures assigned by the American Medical Association which are used throughout the U.S. medical system. Insurance companies decide on which CPT codes can be used for insurance reimbursement.

- Yes
- No or the app has no paid services
- Unable to Assess, please provide a reason: _____

D. Organizational Credibility

Organizational credibility is meant to determine whether the app comes from a trusted source.

13. Who is the developer of the app?

This information can be found on the app website or in the Terms of Service on the app. More than one response may apply. For example, two or more institutions may be involved in the development of the app; or, for example, if an app was developed by Johns Hopkins University, then “Non-profit institution” and “Academic institution” would apply.

- Government
- For-profit company
- Non-profit institution
- Healthcare institution
- Academic institution
- Insurance company
- Independent developer/s
- Unable to assess
- Other, please specify: _____

14. Does the app have any consumer bureau complaints or lawsuits pending?

This information can be found here: Better Business Bureau: <https://www.bbb.org/us/ca/san-francisco/profile/mobile-apps/calmcom-inc-1116-877519/complaints> ; Google Search: <Consumer Complaints> <app name> <app/mobile app>; Google Search: <law suits> <app> <app/mobile app>

- Yes
- No
- Unable to assess

E. Evidence & Clinical Foundation

The information requested in these questions can be found on the app website.

15. Does the app appear to do what it claims to do?

Look at the website to see the claims made by the app and compare them to what you see in the app.

- Yes, the app provides the functionality it claims on its website
- The app provides some of the functionality it claims on its website
- No, the app does not provide the functionality it claims on its website
- Unable to assess

F. Privacy & Security

Security is about the safeguarding of data and unauthorized access of the app. Privacy is about the safeguarding of the user’s identity.

The information requested in these questions should be available in the app in the Terms & Conditions or Privacy & Security agreement. Do not look at information provided on the website.

16. Does the app claim it meets HIPAA [or analogous national standard for protected health information (PHI)]?

- Yes
- No

17. Does the app claim it meets COPPA [or analogous national standard for protected health information (PHI) for minors younger than 13 years of age]?

The Children's Online Privacy Protection Act (COPPA) applies to the online collection of personal information by persons or entities under U.S. jurisdiction about children younger than 13 years of age.

- Yes

- No

18. Does the app report sharing or selling of data for research or commercial purposes?

Information used internally within the app itself to make better recommendations to the user does not count as sharing or selling. Examples of apps that may be sharing or selling data include those that provide evidence based drug and dosage information, clinical decision support systems where you share conditions, symptoms and medication information, medication adherence apps, etc.

- Yes
- No
- Unable to assess

19. If your response to question 18 is “YES”, please answer the following question:

If the app reports sharing or selling for research or commercial purposes, are the data de-identified?

This question is targeted at understanding if the data is de-identified or not. Information used internally, within the app itself, to make better recommendations to the user does not count as sharing or selling.

- Yes, the app reports that data is de-identified
- No, the app does not report data is de-identified
- Unable to assess

20. If the app has the capability to read/write to an electronic health record management system (EHRs) or other healthcare systems, does it use industry standards for secure interoperability (e.g., FHIR, SMART, OAuth 2.0, TLS 1.2)?

This information may need to be deduced. If the app has existing integrations with known EHRs such as EPIC, it can be concluded that it follows industry standards. This information may also need to be identified through a review of the app developer webpage.

- Yes, the app uses industry standards for interoperability
- No, the app does not use industry standards for interoperability
- Not applicable, because it does not read/write to EHRs
- Unable to assess

G. Informed Consent

Informed consent is permission granted by the user after information about the potential use or disclosure of the information collected by the app is provided. Most apps tend to have a disclosure list that is long and hard to understand. There are best practices for ensuring that users understand exactly what they are agreeing to before they “Agree” to the app’s privacy and security practices. The following questions evaluate whether the app follows these best practices.

The information requested in these questions can be found within the app.

21. Please assess the level of informed consent enabled by the app.

Providing consent is often one of the first things users are asked to do after downloading an app. Traditional written consent forms (e.g., Terms of Service, Privacy and Security agreement) are often too long, difficult to understand, and overly focused on legal concerns. It is important that apps enable informed consent for prospective users in a manner that is easy to understand.

- No informed consent
- Does not simplify informed consent
- Average: Focuses on the essential by a) providing a narrative focused on the most salient information, b) limiting concepts to one per screen, and c) following national plain language and health literacy guidance
- Good: Focuses on the essential and organizes content deliberately by a) prioritizing key words/concepts presented on each screen, b) providing information tiers for conceptual elaboration, and c) enabling participants to navigate to their desired level of detail
- Excellent: Focuses on the essential, organizes content deliberately, and encourages engagement through interactive elements

22. What is the format of the data privacy and security consent process followed in the app?

*An opt-in consent process requires **explicit consent from the user** before the collection and processing of their personal data. It refers to a **positive action** taken by the user indicating that they agree to the use of their personal data. An opt-out consent process **does not require the user’s consent** prior to the collection and processing of their personal data. It refers to the process by which users **withdraw their consent** to the use of their personal data.*

- Opt-in for data to be shared, the default choice is opt-out
- Opt-out of data sharing, the default choice is opt-in
- No choice to opt-out of data sharing, without explicit consent you can’t use the app
- Unable to assess

H. Cultural Competence

Cultural competence is defined as the ability to understand, appreciate, and account for different cultures or belief systems based on race, ethnicity, income strata, religious beliefs, etc. It is important to assess if the app captures these differences and whether it provides personalized care that takes these cultural differences into account.

The information requested in these questions can be found within the app or on the app website.

23. Does the app report developing and testing the app for specific cultural group/s?

According to the U.S. Department of Health and Human Services (HHS), "cultural competence is the integration and transformation of knowledge about individuals and groups of people into specific standards, policies, practices, and attitudes used in appropriate cultural settings to increase the quality of services; thereby producing better outcomes." For this assessment, groups with lived experiences such as pregnant teens and survivors of gender-based violence would be considered to be a specific cultural group.

- Yes
- No
- Unable to assess
- Other, please specify: _____

24. If the app reports developing and testing the app for specific cultural groups, is there published documentation (e.g., website, papers) of the process taken to incorporate information that is specific to specified cultures?

Select all that apply:

- Website
- Documentation (on the app or website)
- Published scientific papers
- Unable to assess
- Other, please specify: _____
- Not applicable

If the app integrates culturally specific groups, please name the groups.

25. Is gender-inclusive language employed? When asking the user’s gender, is there an option to self-describe, an option to decline to answer, use of scientifically correct terms for gender (e.g., man, woman, nonbinary)?

This information can be found while assessing the app and is generally encountered when a user creates an account on the app.

Select all that apply:

- An option for use of personal pronouns (e.g., they/them, he/his, she/her)
- An option to specify another gender
- An option to decline to answer
- Use of scientifically correct terms for gender, such as man, woman, non-binary
- Unable to assess
- Not applicable

26. If the app was tested in a study, what was the percentage of non-white participants?

For clinical trials, this information may be found on the clinicaltrials.gov page of the trial under Study Results (include “Other” as non-white). If the study was a usability or other type of study, this information may be found on the developer website.

- Less than 30%
- Between 30 – 50%
- More than 50%
- No disaggregate data available
- No information available

I. Usability

Usability can be described as the capacity of a system to provide a condition for its users to perform the tasks safely, effectively, and efficiently. It is important that the experience be engaging and pleasing, otherwise users are likely to stop using the app.

The information requested in these questions can be found within the app.

27. Does the app work offline?

Choose Yes if parts of the app or the entire app work offline.

- Yes
- No
- Unable to assess

28. Performance: How accurately/fast do the app features (functions) and components (buttons/menus) work?

- App is broken; no/insufficient/inaccurate response (e.g., crashes/bugs/broken features)
- Some functions work, but app lags or contains major technical problems
- App works overall; some technical problems need fixing or app is slow at times
- Mostly functional with minor/negligible problems
- Perfect/timely response; no technical bugs found or contains a 'loading time left' indicator, if relevant

29. What languages are supported by the application?

This information can be found in the App Store.

Select all that apply:

- English
- Spanish
- French
- German
- Other, please specify: _____

30. Customization: Does the app allow customization of settings and preferences by the user (e.g., sound, content, notifications)?

- Does not allow any customization or requires setting to be input every time
- Allows little customization and that limits the app's functions
- Basic customization to function adequately
- Allows numerous options for customization
- Allows complete tailoring to the user's characteristics/preferences, remembers all settings

31. Target group: Is the app content (visuals, language, design) appropriate for the target audience?

- Completely inappropriate, unclear, or confusing
- Mostly inappropriate, unclear, or confusing
- Satisfactory, but not specifically designed for the target audience, may be inappropriate/unclear/confusing at times
- Designed for the target audience, with minor issues
- Designed specifically for the target audience; no issues found

32. Layout: Is the arrangement and size of buttons, icons, menus, and content on the screen appropriate?

- Very bad design: cluttered, some options impossible to select/locate/see/read
- Bad design: random, unclear, some options difficult to select/locate/see/read
- Satisfactory: few problems with selecting/locating/seeing/reading items
- Mostly clear: able to select/locate/see/read items
- Professional: simple, clear, orderly, logically organized

33. Graphics: How high is the quality/resolution of graphics used for buttons, icons, menus, and content?

- Graphics appear amateur, very poor visual design: disproportionate, stylistically inconsistent
- Low quality/low resolution graphics, low quality visual design: disproportionate
- Satisfactory quality graphics and visual design: generally consistent in style
- High quality/resolution graphics and visual design: mostly proportionate, consistent in style
- Very high quality/resolution graphics and visual design: proportionate, consistent in style throughout

34. Visual appeal: How good does the app look?

- Ugly/unpleasant to look at: poorly designed, clashing, mismatched colors
- Bad: poorly designed, bad use of color, visually boring
- Satisfactory: average, neither pleasant nor unpleasant
- Pleasant: seamless graphics, consistent and professionally designed
- Beautiful: very attractive, memorable, stands out; use of color enhances app features/menus

35. Does the app have advertising?

- Yes
- No

36. If app has advertising, is the advertising intrusive and distracting?

- Advertising is neither intrusive nor distracting
- The advertising is both intrusive and distracting
- Not applicable

37. Ease of use: How easy is it to learn how to use the app; how clear are the menu labels, icons, and instructions?

- No/limited instructions, menu labels and icons are confusing, complicated
- Takes a lot of time or effort
- Takes some time or effort
- Easy to learn (or has clear instructions)
- Able to use app immediately, intuitive and simple (no instructions needed)

38. Navigation: Does moving between screens make sense? Does the app have all necessary links between screens?

- No logical connection between screens at all/navigation is difficult
- Understandable after a lot of time/effort
- Understandable after some time/effort
- Easy to understand/navigate
- Perfectly logical, easy, clear and intuitive screen flow throughout, and/or has shortcuts

39. Gestural design: Do taps/swipes/pinches/scrolls make sense? Are they consistent across all components/screens?

- Completely inconsistent/confusing
- Often inconsistent/confusing
- Satisfactory with some inconsistencies/confusing elements
- Mostly consistent/intuitive with negligible problems
- Perfectly consistent and intuitive

40. Content: Is the app copy well written and relevant to the goal and/or topic of the app?

- There is no information within the app
- Irrelevant/inappropriate/incoherent/incorrect
- Barely relevant/appropriate/coherent, may be incorrect
- Satisfactory with respect to relevance/appropriateness/coherence and appears to be correct
- Relevant/appropriate/coherent/correct
- Highly relevant/appropriate/coherent/correct
- Not applicable

41. Is there any evidence to show the average duration of use of the app by users?

This information may be available on the app website or in a paper published in a scientific journal. Do not look at user reviews on the app website, in the App Store, or in news articles.

- On average, users have engaged with the app for 30 days or less
- On average, users have engaged with the app for more than a month but less than 6 months
- On average, users have engaged with the app for more than 180 days
- No evidence available

J. Remote Monitoring

Remote patient monitoring enables monitoring of patients outside of conventional clinical settings through real-time access to patient health data. The patient's provider may have a schedule to monitor patient data or may receive alerts about their patient's health.

The information requested in these questions can be found within the app/app website.

42. If remote monitoring is feasible, how does the provider access the data?

- Provider has access to patient information through the app
- Provider has access to the patient information through the EMR
- Provider has no access to patient information
- Not applicable
- Unable to assess
- Other, please specify: _____

43. Does the app provide alerts to the provider to notify them of a clinical event that may require action on their part?

Two-way messaging capability with a healthcare provider or concierge may be considered to be an alerting mechanism.

- Yes
- No
- Unable to assess

44 . Can the app share data with wearables like Apple Watch and Fitbit?

This option will either pop up on the app during sign-up or be mentioned on the app developer's website.

- Yes
- No
- Unable to assess

K. Access to Crisis Response Services

This section evaluates whether the app provides access to emergency sources of information.

The information requested in this question can be found within the app or on the app website.

45. Does the app connect the user automatically to resources in case of a crisis situation or emergency?

- Yes
- No
- Unable to assess

L. Artificial Intelligence (AI)

Apps often use artificial intelligence (AI) for a variety of tasks including risk-prediction, differential diagnoses, and personalization of health content. The following questions are important because we want to gauge the potential of the app to cause harm and also to determine whether the app's algorithms are being updated based on user input.

The information requested in these questions can be found in the app, on the app website, or from evidence found through a literature search.

46. Does the app use AI? If yes, please respond to questions 47 and 48.

- Yes
- No

47. How is AI being used in the app?

For example, AI may be used in the app to automate simple tasks, such as scheduling, or it may be used for more complex functionality, such as to personalize chatbots.

48. Is there any evidence to suggest that the app uses data from user interactions to improve precision on AI models?

For example, after an interaction, the app may ask the user for feedback on the automated response from the app to see if it was appropriate. If the response was not correct, the app may ask for further feedback from the user so the algorithms can be refined.

- Yes, there is evidence to suggest that the AI models are updated based on feedback
- No, there is no evidence to suggest that the AI models are updated based on feedback
- Unable to assess

Section 3: Mental Health App Features

This section attempts to capture the features specific to mental health present in the app. This is not meant to be an exhaustive list. Only some of these features may apply to the app being reviewed.

The information requested in these questions can be found in the app, on the app website or in published literature.

1. List mental health symptom(s) and/or condition(s) addressed by the app.

bipolar disorder

2. Please answer the following questions about the app features.

Question	Yes	No	Unable to assess
a. Does the app facilitate text messaging interactions with healthcare therapists, coaches, or other providers via the app?		x	
b. Does the app facilitate audio chat interactions with healthcare therapists, coaches, or other providers via the app?		x	
c. Does the app facilitate video chat interactions with healthcare therapists, coaches, or other providers via the app?		x	
d. Does the app facilitate teletherapy services via the app?		x	
e. Does the app facilitate group therapy services via the app?		x	
f. Does the app provide live support to a coach or counselor via the app?		x	
g. Does the app provide concierge mental health services via the app? Concierge services are personalized services patients can choose based on what they think works best for them. Some apps may provide acute concierge services to include additional help during times of crisis or high stress.		x	

3. Does the app provide a direct connection to 988 or other hotlines?

- Yes, and the connection works
- Yes, but the connection does not work
- No
- Unable to assess

4. Are the following functionalities supported by the app?

The following questions ask you to rate the functionality as either “comprehensive” or “not comprehensive.”

Comprehensive: Broad range of functionality that meets or exceeds the needs of the user and allows for flexibility

Not comprehensive: Rudimentary functionality that may not meet the needs of the user or provides limited choices and lacks flexibility

	Comprehensive	Not Comprehensive	Not Applicable
Mindfulness: Mindfulness is a therapeutic technique that includes elements of relaxation, breathing, and body exercise. It also includes techniques, such as meditation, guided positive imagery, grounding exercises, or progressive muscle relaxation.			x
Journaling: Journaling can help users manage anxiety, reduce stress, and cope with mental health challenges. It can provide users with a way of identifying negative thoughts and behaviors and highlight positive aspects in their lives.			x
Psychoeducation: Psychoeducation may be defined as the education of a person with a psychiatric disorder regarding the symptoms, treatments, and prognosis of that illness. Psychoeducation may also be targeted to the caregiver, family member, or loved one of the patient. Psychoeducation should be a brief personal intervention by a healthcare provider upon first diagnosis; however, sometimes, only written materials or online resources are provided to the patient. In the worst case scenario, patients receive no Psychoeducation from their provider. (Source: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD010823.pub2/full)		x	
Skill building: Skill building includes techniques for recognition of signs and symptoms of a problem, self-advocacy, stress management, emotional regulation, relapse prevention, promotion of sleep hygiene, etc. The app may include tips and advice on dealing with negative emotions and with behavior change. Skill building may require repetition, positive reinforcement, modeling, and practice.			x
Screening: Historically, mental health care has relied on structured patient interviews and self-reported questionnaires for diagnosis. These can be used for self-evaluation, reporting, or to decide whether a patient should engage with a mental health professional. Please indicate if industry-validated screening questionnaires (e.g., GATT, PHQ-4, PHQ-9) are provided by the app. The standardized screening questionnaires used will generally be mentioned on the app developer's website.			x
Safety planning: Safety planning is designed to help individuals respond to escalation of suicidal thoughts and feelings, giving them a tailored list of coping behaviors and social support to use until the risk passes.			x

	Comprehensive	Not Comprehensive	Not Applicable
Sleep Hygiene: This includes features for promoting sleep hygiene, such as sleep diaries, lifestyle tracking, alarms, data synchronization with a wearable device (e.g., FitBit, Apple Watch), etc.			x
Automated Chatbots: Chatbots include conversational agents that can provide virtual therapy and/or social support.			x
Family/Caregiver Support: Caregivers take care of the day-to-day needs of patients. Apps can help caregivers monitor the mental state of the patient, identify the early signs of illness, track relapse and deterioration, and help the patient access services. The family/caregiver may also supervise treatment and provide emotional support to the patient.			x
Social and peer group interaction: Individuals with mental health conditions may find it hard to engage socially and may have a small social network/peer group. Apps may provide peer group interactions and/or community support discussion boards for such individuals.			x
Gamification: Gamification is the application of typical elements of game playing (e.g., point scoring, competition with others, rules of play) to other areas of activity. Gamification is used to encourage and reward positive changes.			x
Personalization: Personalized treatment entails the selection of a therapy or treatment protocol based on a patient's profile, which may increase the likelihood of a successful outcome. Personalization can be used to tailor interactions initially, when the user starts to use the app, and also on a more regular basis. It is usually based on user information collected by the app, such as what activities produce positive emotions (e.g., physical exercise, talking to a friend).			x
Self-Monitoring: Self-monitoring apps allow users to engage in symptom monitoring, mood tracking, and progress tracking through a treatment program, such as one for substance abuse, depression and anxiety, etc.			x
Medication Adherence: Does the app support patients in their use of medication(s)? This could be through alerts, notifications, or other means.			x
Medication Delivery: Does the app support physical delivery of medications to patient(s)?			x

5. Please add any features supported by the app that are not listed above.

n/a

Post Administrative Questionnaire

This section contains some details about the app and evidence being reviewed.

1. Include all links to references for evidence (citations on PubMed, systematic reviews, websites, etc.) here.

n/a

2. Based on your review of the app, do you think it would have been useful to have some training or a tutorial about how to use it?

This is your subjective assessment based on your usage of the app.

- Yes, there was training available, and I needed it to use the app
- Yes, there was training available, but I didn't need it to use the app
- No, there was no training available, but I needed it to use the app
- No, there was no training available, and I didn't need it to use the app
- Don't know

3. Does the app have any help-related documentation available in the app itself (e.g., tooltips, general help)?

Base your answer on what was available on the app itself.

In-app help usually comes in three forms: 1) a separate page or pages of help within the app, with instructions; 2) popups that provide contextual help, displaying instructions relevant to the specific task that the user is attempting; 3) descriptions of app features of interest to the user. If at least two of these features is present, the help-related documentation would be considered to be comprehensive.

- Yes, there was comprehensive help available on the app
- Yes, there was some help available on the app
- The app help button took me to the website
- No help was available on the app

4. Does the app have a dedicated website that provides information about the app?

This may be part of a company website or a standalone website focused solely on the app.

- Yes
- No
- Unable to assess

5. Does the app have any help-related documentation available on its website?

- Yes, there was comprehensive help available
- Yes, there was some help available
- No, help was not available on the app

6. Please provide a subjective evaluation of the app in your own words.

- Do you think the assessment conducted here matches your subjective assessment? If not, please provide additional details.
- Do you think the risk posed by the app is captured accurately here?
- Do you think the app is technically sound?
- Do you think the therapeutic content provided by the app is rudimentary or substantive?
- Can this app help those that may be experiencing mental or behavioral health challenges?

very poor quality, information was written poorly and likely inaccurate/incomplete

7. Do you think the app could cause harm or have a negative impact on the user? Please elaborate.

Perhaps if the user used this app as their source of truth about bipolar disorder.

8. Do you think the subjective risk assessment matches the risk assessment calculated in Section 1? Please elaborate.

Yes, mostly (see question 7)

9. Any additional comments?

n/a

10. Please note the time taken to complete your assessment using this framework.

45 minutes

Assessment Example, App B

Administrative Questionnaire

This section contains introductory questions about the app to be reviewed. The questions are mostly factual, rather than evaluative. Some of these questions may make it into the final Assessment Framework, based on the information we gather from the pilot.

Q 1. Reviewer Name

Q 2. Date of Evaluation

Q 3. App Name

Q 4. App/ Company Website

This should only be a specific website developed by the app developer and not their pages on social media sites, such as Instagram or Facebook.

Q 5. Country of Origin

This information can be found in the Terms & Conditions on the website or in the app.

Q 6. App Version Number

This information can be found on the app and from the App Store.

Q 7. Platform used for assessment by the reviewer

Q 8. OS Version Number on Evaluating Device

*For Android go to: Settings > About Phone > Software Information;
For iOS go to: Settings > General > Software Version.*

Q 9. Describe the goal of the app in your own words. Do not use any of the marketing phrases/lingo used by the app developer.

This should be based on the app description on the website.

A journaling-based app for self care and mindfulness with free text journaling and mood tracking

Q 10. Has the app been approved by any regulating authority, such as FDA?

This information can be found on the app developer website.

No

Q 11. Does the app have multiple revenue models, such as freemium, in-app purchases, etc.?

- Yes
- No

Q 12. If the app has multiple versions (e.g., free, freemium, paid), which version did you evaluate? If the app had only one revenue model, please choose “not applicable” here.

- Free
- Paid version
- Paid by insurance agency, employer, or healthcare provider
- Not applicable
- Unable to assess

Q 13. Does the app require prior authorization from a healthcare institution/insurance provider/college/employer or other institution for access?

- Yes
- No
- Unable to assess

Q 14. If you were unable to assess the app, please specify the reason.

Able to access

Note: Questions about privacy, security, and informed consent are covered in both Section 1 and Section 2 of the framework. Please make a note of all the disclaimers, warnings, and privacy and security statements and agreements when you first log into the app, so you are able to answer the questions that come later in the framework.

Section 1: Risks and Mitigation Strategies

Section 1 aims to evaluate the potential risk posed by the app, and evidence of efficacy, and safety features. The burden of evidence to prove efficacy and safety is dependent on the goals of the app, the target audience, and the severity of the mental health condition.

The categories of questions in this section are App Integrity, Risk Assessment, Evidence, Linkage to Care, and Access to Crisis Services. If apps do not meet the threshold for safety and credibility, that is flagged in the assessment.

A. App Integrity

The information requested in these questions can be found on the App Store, the app's website, or via Google Search.

1. Was the version of the app you are reviewing updated in the last 6 months?

- Yes
- No
- Unable to assess

2. Does the app provide a privacy and security agreement to be reviewed and agreed to by the user?

The agreement should be accessible through a link on the app or on the mobile app. Reviewers should not look for this information in other locations, such as the developer's website.

- Yes, there is a privacy and security agreement to be agreed to by the user
- No, there is not a privacy and security agreement
- Unable to assess or unable to access

3. Does the app provide warnings and disclaimers (e.g., the limitations of the app related to medical liability)?

Reviewers should look for this information on the mobile app ONLY.

- Yes, warnings and/or disclaimers are provided
- No, warnings and/or disclaimers are not provided
- Unable to assess
- Not applicable

4. Has the app been endorsed by, or is it being used by, a government agency or trusted mental health professional association?

Endorsement and/or usage helps us establish the credibility of the app.

Endorsement means to give one's approval to, especially officially by an organization. An app merely being reviewed by the American Psychological Association (APA), or a similar organization, does not mean it has been approved by them. Endorsement should be for the app and not the company developing the app or the domain (e.g., teletherapy). Self-reported endorsements are acceptable.

Usage by an institution signifies that an app is made available to all its members. For example, Veterans Affairs uses iBlueButton to deliver critical health care information to veterans.

Examples of government agencies include federal, state and local agencies such as CMS, Veterans Affairs, Virginia Department of Health. Examples of trusted mental health professional associations include Substance Abuse and Mental Health Services Administration (SAMHSA), American Psychological Association (APA), and National Alliance on Mental Illness (NAMI).

- Yes, the app has been endorsed by one or more mental health associations, government agencies, or non-government bodies (examples may include but are not limited to government agencies, such as Centers for Medicare & Medicaid Services (CMS) or Veterans Affairs (VA); non-government bodies, such as APA; and insurance or healthcare institutions)
- No, the app has not been endorsed
- Unable to assess, please specify reason: _____

App Integrity Assessment

Question	Response	App Integrity Level
Does the app satisfy the Integrity Assessment?	<input type="radio"/> Responses to questions 1-3 are all Yes OR response to question 4 is Yes	App Integrity is "High"
Please pick only 1 response.	<input checked="" type="radio"/> Response to any of the questions 1-3 is No AND response to 4 is No	App Integrity is "Low"

B. Risk Assessment

This category aims to determine the level of risk posed by the app. For example, if the app is used by someone with a clinically diagnosed condition, such as Schizophrenia, then the risk posed by the app may be higher than if the person has mild anxiety. Similarly, if the app targets children and adolescents, then the potential risk associated with the app may be higher than if it targeted adults with no impairment or mental health diagnosis.

*For the purposes of this report, either minors (younger than 18 years of age) or adults with a moderate or severe level of functional impairment that causes them to need support from a caregiver are considered to be part of a **vulnerable population**.*

The information requested in these questions can be found within the app/app website.

5. Is the app intended for use by adults who may have an illness or disability that impacts their decisionmaking ability? Severe autism or severe dementia are examples of illnesses or disabilities that may affect decisionmaking ability.

For the purposes of this assessment, the population that may have an illness or disability that impacts their decision making ability is included in the "vulnerable population" group. Adults can usually make medical decisions, such as consenting to treatment, on their own. In some cases, adults need someone else to make medical decisions for them if their decisionmaking ability is affected by an illness or disability. This is sometimes referred to as "substituted consent." A higher level of evidence/protection is needed for the use of specific apps under these circumstances.

- Yes
- No
- Unable to assess

6. Does the app claim that it is intended for use by minors (i.e., people younger than 18 years of age)?

For the purposes of this assessment, this population is included in the "vulnerable population" group. The app is "intended for use" by a population if it is clearly specified on the app or website that this population is its target user group. Do not look at the App Store Age Ratings or Terms and Conditions for this information.

- Yes, for use by minors
- Yes, for use by both minors and adults
- No, not for use by minors
- Unable to assess

7. Does the app claim to provide standalone treatment for any mental health condition?

A standalone treatment might be some form of psychotherapy or other any other medical care provided by the app independently, rather than in conjunction with a licensed healthcare practitioner.

Functionality, such as mindfulness training, skills training, symptom or mood tracking, and social support, would not be considered standalone treatment.

- Yes, the app provides standalone treatment

- No, the app does not provide standalone treatment
- Unable to assess

C. Evidence

It is important that the app has a solid clinical foundation. If the app targets a vulnerable population, the burden of evidence to prove efficacy and safety is greater.

The information requested in these questions can be found on the app website (look at customer testimonials or resources/blog for anecdotal information) and on PubMed (search PubMed as outlined in the training document). Make sure to save all the references in the app review document.

8. Has the app been evaluated for efficacy/effectiveness through a scientifically validated study?

This can be determined by evaluating whether a research study about the app has been published in a peer-reviewed journal. The research should be specific to the mental health component of the app.

- Yes, there is strong research support for the app (i.e., at least one published paper in a peer-reviewed journal that uses a randomized trial that shows efficacy or effectiveness)
- Yes, there is some research support for the app (i.e., at least one published paper in a peer-reviewed journal that uses single-case design, quasi-experimental methods demonstrating efficacy)
- No, there is no research support for the app

9. Is the app based on or does it use an evidence-based strategy, such as Cognitive Behavioral Therapy (CBT), Dialectical Behavior Therapy (DBT), or evidence-based guidelines?

Evidence-based medicine aims to assess the strength of proof behind medical interventions in terms of risks and benefits and, therefore, can be used to inform clinical decisionmaking on both an individual and a population basis.

- Yes, the app reports using an evidence-based strategy/guidelines to achieve its goals
- No, the app does not report the use of evidence-based strategy/guidelines to achieve its goals (or there are no goals described)
- Unable to assess

D. Linkage to Care

This category evaluates the linkages to healthcare providers, caregivers. This information can be found within the app/app website.

10. Does the app facilitate remote monitoring of the patient or send alerts to a clinician/clinical care team or caregiver?

This question seeks to assess whether a healthcare provider or caregiver can monitor the user's health data in real time AND/OR receive alerts triggered by certain pre-specified high-risk events. Functionality to export data into a file or an email does not count as facilitation of remote monitoring.

- Yes
- No
- Unable to assess

10.a. If YES, please specify *who* can monitor the health of the patient, either through alerts or by other means?

- Healthcare provider only
- Caregiver only
- Both healthcare provider and caregiver
- Other
- Not applicable

E. Access to Crisis Services

This section evaluates whether the app provides access to emergency sources of information.

This information can be found within app/app website. There should be a caregiver or healthcare provider interface within the app. Exporting data into a file or an email would not be considered facilitation of remote monitoring.

11. Does the app provide users with information about resources that can be reached in case of an emergency? This information should be available on the mobile app. Reviewers should not look for this information on the website.

- Yes (e.g., a crisis hotline, 911, nearest emergency room services.)
- No, the app provides no information on emergency services
- Unable to assess

12. Evidence and Linkage to Care: If the app can be used by individuals that require substituted consent OR by minors, then is consent sought from either a caregiver/parent/legal guardian?

This should be assessed at the time of creating a login for the app.

- Yes
- No
- Unable to assess
- Not applicable

Risk Assessment

What is the Risk Category for the App?

A lower number signifies a lower risk level.

Risk Level 1: Apps rated as Risk Level 1 do not target a vulnerable population AND do not provide standalone treatment. These are likely to be apps that provide functions such as mindfulness, meditation, and wellness resources. No specific safety checks are needed for these apps.

- Risk Level is 3:** Response to EITHER Question 5 or Question 6 is Yes AND response to Question 7 is Yes, *go to Safety Assessment Step 3*
- Risk Level is 2:** Response to EITHER Question 5 or Question 6 is Yes OR response to Question 7 is Yes, *go to Safety Assessment Step 3*
- Risk Level is 1:** None of the above, *go to Safety Assessment Step 3*

Safety Assessment

Step 1: Preliminary Safety Assessment for Risk Levels 2 & 3

Risk Level 2	<input type="radio"/> [If responses to Questions 9-11 are Yes AND response to Question 10 is Yes or No, supporting evidence-base available]	PRELIM SECURITY CHECK PASSED , <i>go to Safety Assessment Step 2</i>
	Else	PRELIM SECURITY CHECK FAILED , <i>go to Safety Assessment Step 2</i>
Risk Level 3	<input type="radio"/> [If responses to Questions 8-11 are Yes, supporting evidence-base available]	PRELIM SECURITY CHECK PASSED , <i>go to Safety Assessment Step 2</i>
	Else	PRELIM SECURITY CHECK FAILED , <i>go to Safety Assessment Step 2</i>

Step 2: Safety Assessment for Vulnerable Population

If response to Question 5 and/or Question 6 is Yes,

- If response to Question 12 is **Yes**, then **SAFE** for vulnerable populations, go to *Safety Assessment Step 3*
- If response to Question 12 is **No**, then **NOT SAFE** for vulnerable populations, go to *Safety Assessment Step 3*

else go to *Step 3 under Safety Assessment*

STEP 3: Safety Assessment

RISK LEVEL	PRELIM SAFETY CHECK	VULNERABLE POPULATION	APP SAFETY
Risk Level 1	N/A	N/A or SAFE for Vulnerable Populations	Passed Safety Check
	N/A	Not Safe for Vulnerable Populations	Failed Safety Check
Risk Level 2	PASS	N/A or SAFE for Vulnerable Populations	Passed Safety Check
		Not Safe for Vulnerable Populations	Failed Safety Check
	FAIL	--	Failed Safety Check
Risk Level 3	PASS	N/A or SAFE for Vulnerable Populations	Passed Safety Check
		Not Safe for Vulnerable Populations	Failed Safety Check
	FAIL	--	Failed Safety Check

Section 2: Function

This section is focused on descriptive aspects, such as accessibility, usability, privacy/security, AI, cultural competence, and informed consent.

A. Accessibility Features

Accessibility features are meant to make the app easier to navigate, especially for those with disabilities. Common accessibility features include text-to-speech, closed-captioning, and keyboard shortcuts.

1. Which of the phone's accessibility features work within the app?

Reviewers will have to activate accessibility features on their phone under Settings in order to test this.

Select all that apply:

- Text adjustment feature
- Colorblind color scheme feature
- Text-to-speech feature
- None of the phone's accessibility features work in the app
- Unable to assess, please provide a reason: _____

2. Are there additional accessibility features that are provided by the app?

These are accessibility features that are provided within the app but cannot be set as an option under Settings.

Select all that apply:

- Adaptation of audio/video content with transcriptions or captions
- Tapping and other gestures are configurable
- Contrast text coloring in the content
- Screen reader
- No additional accessibility features provided by the app
- Other features not provided here (please specify): _____
- Unable to assess, please provide a reason: _____

B. App Information

General info about the app and user ratings.

The information requested in these questions can be obtained by going to the appropriate app store using your browser.

3. Does the app work on Apple(iOS)?

- Yes
- No

4. What are the number of reviews on the iOS App Store?

The iOS App Store can be accessed through a browser on your phone, tablet, or computer.

78400

5. What is the app rating (number of stars) on Apple Store?

This provides users' perspectives of the app.

4.6

6. Does the app work on Android?

- Yes
- No

7. What are the number of reviews on Google Play Store.

38932

8. What is the app rating (number of stars) on Google Play Store?

The Google Play Store can be accessed through a browser on your phone, tablet, or computer.

4.1

C. Costs

App costs could be upfront, in the form of a monthly/annual subscription, or through freemium services that require in-app purchasing. Some apps require payment from the user, while others are reimbursed by the healthcare provider or insurance.

The information requested in these questions may be found in the app, in the App Store, and on the app site.

9. What is the business model for the app? More than one option may be applicable here. An app that provides a free trial for a limited duration is not considered to be free.

If the app provides a payment gateway to pay a therapist or healthcare provider, those costs should not be included.

- Free (no upfront fee for the app; additional packages/services may be offered for a fee)
- Upfront fee (a onetime cost for accessing the app; additional packages/services may be offered for additional fees)
- In-app purchases (additional packages/services available for a fee in addition to or in lieu of an upfront cost, e.g., concierge services)
- Subscription (payment for services on a monthly/quarterly/annual basis)
- Reimbursed by healthcare providers/insurers/employers
- Other, please specify: _____

10. Does the app provide a free or freemium model? An app that provides a free trial for a limited duration is not considered to be free.

- Free
- Freemium
- No free or freemium version

11. What is the estimated annual cost of the app for the paid version?

If the app provides a free or freemium model, then select “Not applicable” in the list below.

- Under \$50
- \$50-250
- Over \$250
- Not applicable

12. If the app includes paid service(s), does it provide CPT code(s) for insurance reimbursement?

Current Procedural Terminology (CPT) are unique 5-digit codes for medical services or procedures assigned by the American Medical Association which are used throughout the U.S. medical system. Insurance companies decide on which CPT codes can be used for insurance reimbursement.

- Yes
- No or the app has no paid services
- Unable to Assess, please provide a reason: _____

D. Organizational Credibility

Organizational credibility is meant to determine whether the app comes from a trusted source.

13. Who is the developer of the app?

This information can be found on the app website or in the Terms of Service on the app. More than one response may apply. For example, two or more institutions may be involved in the development of the app; or, for example, if an app was developed by Johns Hopkins University, then “Non-profit institution” and “Academic institution” would apply.

- Government
- For-profit company
- Non-profit institution
- Healthcare institution
- Academic institution
- Insurance company
- Independent developer/s
- Unable to assess
- Other, please specify: _____

14. Does the app have any consumer bureau complaints or lawsuits pending?

This information can be found here: Better Business Bureau: <https://www.bbb.org/us/ca/san-francisco/profile/mobile-apps/calmcom-inc-1116-877519/complaints> ; Google Search: <Consumer Complaints> <app name> <app/mobile app>; Google Search: <law suits> <app> <app/mobile app>

- Yes
- No
- Unable to assess

E. Evidence & Clinical Foundation

The information requested in these questions can be found on the app website.

15. Does the app appear to do what it claims to do?

Look at the website to see the claims made by the app and compare them to what you see in the app.

- Yes, the app provides the functionality it claims on its website
- The app provides some of the functionality it claims on its website
- No, the app does not provide the functionality it claims on its website
- Unable to assess

F. Privacy & Security

Security is about the safeguarding of data and unauthorized access of the app. Privacy is about the safeguarding of the user's identity.

The information requested in these questions should be available in the app in the Terms & Conditions or Privacy & Security agreement. Do not look at information provided on the website.

16. Does the app claim it meets HIPAA [or analogous national standard for protected health information (PHI)]?

- Yes
- No

17. Does the app claim it meets COPPA [or analogous national standard for protected health information (PHI) for minors younger than 13 years of age]?

The Children's Online Privacy Protection Act (COPPA) applies to the online collection of personal information by persons or entities under U.S. jurisdiction about children younger than 13 years of age.

- Yes
- No

18. Does the app report sharing or selling of data for research or commercial purposes?

Information used internally within the app itself to make better recommendations to the user does not count as sharing or selling. Examples of apps that may be sharing or selling data include those that provide evidence based drug and dosage information, clinical decision support systems where you share conditions, symptoms and medication information, medication adherence apps, etc.

- Yes
- No
- Unable to assess

19. If your response to question 18 is “YES”, please answer the following question:

If the app reports sharing or selling for research or commercial purposes, are the data de-identified?

This question is targeted at understanding if the data is de-identified or not. Information used internally, within the app itself, to make better recommendations to the user does not count as sharing or selling.

- Yes, the app reports that data is de-identified
- No, the app does not report data is de-identified
- Unable to assess

20. If the app has the capability to read/write to an electronic health record management system (EHRs) or other healthcare systems, does it use industry standards for secure interoperability (e.g., FHIR, SMART, OAuth 2.0, TLS 1.2)?

This information may need to be deduced. If the app has existing integrations with known EHRs such as EPIC, it can be concluded that it follows industry standards. This information may also need to be identified through a review of the app developer webpage.

- Yes, the app uses industry standards for interoperability
- No, the app does not use industry standards for interoperability
- Not applicable, because it does not read/write to EHRs
- Unable to assess

G. Informed Consent

Informed consent is permission granted by the user after information about the potential use or disclosure of the information collected by the app is provided. Most apps tend to have a disclosure list that is long and hard to understand. There are best practices for ensuring that users understand exactly what they are agreeing to before they “Agree” to the app’s privacy and security practices. The following questions evaluate whether the app follows these best practices.

The information requested in these questions can be found within the app.

21. Please assess the level of informed consent enabled by the app.

Providing consent is often one of the first things users are asked to do after downloading an app. Traditional written consent forms (e.g., Terms of Service, Privacy and Security agreement) are often too long, difficult to understand, and overly focused on legal concerns. It is important that apps enable informed consent for prospective users in a manner that is easy to understand.

- No informed consent
- Does not simplify informed consent
- Average: Focuses on the essential by a) providing a narrative focused on the most salient information, b) limiting concepts to one per screen, and c) following national plain language and health literacy guidance
- Good: Focuses on the essential and organizes content deliberately by a) prioritizing key words/concepts presented on each screen, b) providing information tiers for conceptual elaboration, and c) enabling participants to navigate to their desired level of detail
- Excellent: Focuses on the essential, organizes content deliberately, and encourages engagement through interactive elements

22. What is the format of the data privacy and security consent process followed in the app?

*An opt-in consent process requires **explicit consent from the user** before the collection and processing of their personal data. It refers to a **positive action** taken by the user indicating that they agree to the use of their personal data. An opt-out consent process **does not require the user’s consent** prior to the collection and processing of their personal data. It refers to the process by which users **withdraw their consent** to the use of their personal data.*

- Opt-in for data to be shared, the default choice is opt-out
- Opt-out of data sharing, the default choice is opt-in
- No choice to opt-out of data sharing, without explicit consent you can’t use the app
- Unable to assess

H. Cultural Competence

Cultural competence is defined as the ability to understand, appreciate, and account for different cultures or belief systems based on race, ethnicity, income strata, religious beliefs, etc. It is important to assess if the app captures these differences and whether it provides personalized care that takes these cultural differences into account.

The information requested in these questions can be found within the app or on the app website.

23. Does the app report developing and testing the app for specific cultural group/s?

According to the U.S. Department of Health and Human Services (HHS), "cultural competence is the integration and transformation of knowledge about individuals and groups of people into specific standards, policies, practices, and attitudes used in appropriate cultural settings to increase the quality of services; thereby producing better outcomes." For this assessment, groups with lived experiences such as pregnant teens and survivors of gender-based violence would be considered to be a specific cultural group.

- Yes
- No
- Unable to assess
- Other, please specify: _____

24. If the app reports developing and testing the app for specific cultural groups, is there published documentation (e.g., website, papers) of the process taken to incorporate information that is specific to specified cultures?

Select all that apply:

- Website
- Documentation (on the app or website)
- Published scientific papers
- Unable to assess
- Other, please specify: _____
- Not applicable

If the app integrates culturally specific groups, please name the groups.

25. Is gender-inclusive language employed? When asking the user’s gender, is there an option to self-describe, an option to decline to answer, use of scientifically correct terms for gender (e.g., man, woman, nonbinary)?

This information can be found while assessing the app and is generally encountered when a user creates an account on the app.

Select all that apply:

- An option for use of personal pronouns (e.g., they/them, he/his, she/her)
- An option to specify another gender
- An option to decline to answer
- Use of scientifically correct terms for gender, such as man, woman, non-binary
- Unable to assess
- Not applicable

26. If the app was tested in a study, what was the percentage of non-white participants?

For clinical trials, this information may be found on the clinicaltrials.gov page of the trial under Study Results (include “Other” as non-white). If the study was a usability or other type of study, this information may be found on the developer website.

- Less than 30%
- Between 30 – 50%
- More than 50%
- No disaggregate data available
- No information available

I. Usability

Usability can be described as the capacity of a system to provide a condition for its users to perform the tasks safely, effectively, and efficiently. It is important that the experience be engaging and pleasing, otherwise users are likely to stop using the app.

The information requested in these questions can be found within the app.

27. Does the app work offline?

Choose Yes if parts of the app or the entire app work offline.

- Yes
- No
- Unable to assess

28. Performance: How accurately/fast do the app features (functions) and components (buttons/menus) work?

- App is broken; no/insufficient/inaccurate response (e.g., crashes/bugs/broken features)
- Some functions work, but app lags or contains major technical problems
- App works overall; some technical problems need fixing or app is slow at times
- Mostly functional with minor/negligible problems
- Perfect/timely response; no technical bugs found or contains a 'loading time left' indicator, if relevant

29. What languages are supported by the application?

This information can be found in the App Store.

Select all that apply:

- English
- Spanish
- French
- German
- Other, please specify: _____

30. Customization: Does the app allow customization of settings and preferences by the user (e.g., sound, content, notifications)?

- Does not allow any customization or requires setting to be input every time
- Allows little customization and that limits the app's functions
- Basic customization to function adequately
- Allows numerous options for customization
- Allows complete tailoring to the user's characteristics/preferences, remembers all settings

31. Target group: Is the app content (visuals, language, design) appropriate for the target audience?

- Completely inappropriate, unclear, or confusing
- Mostly inappropriate, unclear, or confusing
- Satisfactory, but not specifically designed for the target audience, may be inappropriate/unclear/confusing at times
- Designed for the target audience, with minor issues
- Designed specifically for the target audience; no issues found

32. Layout: Is the arrangement and size of buttons, icons, menus, and content on the screen appropriate?

- Very bad design: cluttered, some options impossible to select/locate/see/read
- Bad design: random, unclear, some options difficult to select/locate/see/read
- Satisfactory: few problems with selecting/locating/seeing/reading items
- Mostly clear: able to select/locate/see/read items
- Professional: simple, clear, orderly, logically organized

33. Graphics: How high is the quality/resolution of graphics used for buttons, icons, menus, and content?

- Graphics appear amateur, very poor visual design: disproportionate, stylistically inconsistent
- Low quality/low resolution graphics, low quality visual design: disproportionate
- Satisfactory quality graphics and visual design: generally consistent in style
- High quality/resolution graphics and visual design: mostly proportionate, consistent in style
- Very high quality/resolution graphics and visual design: proportionate, consistent in style throughout

34. Visual appeal: How good does the app look?

- Ugly/unpleasant to look at: poorly designed, clashing, mismatched colors
- Bad: poorly designed, bad use of color, visually boring
- Satisfactory: average, neither pleasant nor unpleasant
- Pleasant: seamless graphics, consistent and professionally designed
- Beautiful: very attractive, memorable, stands out; use of color enhances app features/menus

35. Does the app have advertising?

- Yes
- No

36. If app has advertising, is the advertising intrusive and distracting?

- Advertising is neither intrusive nor distracting
- The advertising is both intrusive and distracting
- Not applicable

37. Ease of use: How easy is it to learn how to use the app; how clear are the menu labels, icons, and instructions?

- No/limited instructions, menu labels and icons are confusing, complicated
- Takes a lot of time or effort
- Takes some time or effort
- Easy to learn (or has clear instructions)
- Able to use app immediately, intuitive and simple (no instructions needed)

38. Navigation: Does moving between screens make sense? Does the app have all necessary links between screens?

- No logical connection between screens at all/navigation is difficult
- Understandable after a lot of time/effort
- Understandable after some time/effort
- Easy to understand/navigate
- Perfectly logical, easy, clear and intuitive screen flow throughout, and/or has shortcuts

39. Gestural design: Do taps/swipes/pinches/scrolls make sense? Are they consistent across all components/screens?

- Completely inconsistent/confusing
- Often inconsistent/confusing
- Satisfactory with some inconsistencies/confusing elements
- Mostly consistent/intuitive with negligible problems
- Perfectly consistent and intuitive

40. Content: Is the app copy well written and relevant to the goal and/or topic of the app?

- There is no information within the app
- Irrelevant/inappropriate/incoherent/incorrect
- Barely relevant/appropriate/coherent, may be incorrect
- Satisfactory with respect to relevance/appropriateness/coherence and appears to be correct
- Relevant/appropriate/coherent/correct
- Highly relevant/appropriate/coherent/correct
- Not applicable

41. Is there any evidence to show the average duration of use of the app by users?

This information may be available on the app website or in a paper published in a scientific journal. Do not look at user reviews on the app website, in the App Store, or in news articles.

- On average, users have engaged with the app for 30 days or less
- On average, users have engaged with the app for more than a month but less than 6 months
- On average, users have engaged with the app for more than 180 days
- No evidence available

J. Remote Monitoring

Remote patient monitoring enables monitoring of patients outside of conventional clinical settings through real-time access to patient health data. The patient's provider may have a schedule to monitor patient data or may receive alerts about their patient's health.

The information requested in these questions can be found within the app/app website.

42. If remote monitoring is feasible, how does the provider access the data?

- Provider has access to patient information through the app
- Provider has access to the patient information through the EMR
- Provider has no access to patient information
- Not applicable
- Unable to assess

Other, please specify: _____

43. Does the app provide alerts to the provider to notify them of a clinical event that may require action on their part?

Two-way messaging capability with a healthcare provider or concierge may be considered to be an alerting mechanism.

- Yes
- No
- Unable to assess

44 . Can the app share data with wearables like Apple Watch and Fitbit?

This option will either pop up on the app during sign-up or be mentioned on the app developer's website.

- Yes
- No
- Unable to assess

K. Access to Crisis Response Services

This section evaluates whether the app provides access to emergency sources of information.

The information requested in this question can be found within the app or on the app website.

45. Does the app connect the user automatically to resources in case of a crisis situation or emergency?

- Yes
- No
- Unable to assess

L. Artificial Intelligence (AI)

Apps often use artificial intelligence (AI) for a variety of tasks including risk-prediction, differential diagnoses, and personalization of health content. The following questions are important because we want to gauge the potential of the app to cause harm and also to determine whether the app's algorithms are being updated based on user input.

The information requested in these questions can be found in the app, on the app website, or from evidence found through a literature search.

46. Does the app use AI? If yes, please respond to questions 47 and 48.

- Yes
- No

47. How is AI being used in the app?

For example, AI may be used in the app to automate simple tasks, such as scheduling, or it may be used for more complex functionality, such as to personalize chatbots.

48. Is there any evidence to suggest that the app uses data from user interactions to improve precision on AI models?

For example, after an interaction, the app may ask the user for feedback on the automated response from the app to see if it was appropriate. If the response was not correct, the app may ask for further feedback from the user so the algorithms can be refined.

- Yes, there is evidence to suggest that the AI models are updated based on feedback
- No, there is no evidence to suggest that the AI models are updated based on feedback
- Unable to assess

Section 3: Mental Health App Features

This section attempts to capture the features specific to mental health present in the app. This is not meant to be an exhaustive list. Only some of these features may apply to the app being reviewed.

The information requested in these questions can be found in the app, on the app website or in published literature.

1. List mental health symptom(s) and/or condition(s) addressed by the app.

anxiety, depression, eating disorder, PTSD, SUD

2. Please answer the following questions about the app features.

Question	Yes	No	Unable to assess
a. Does the app facilitate text messaging interactions with healthcare therapists, coaches, or other providers via the app?		X	
b. Does the app facilitate audio chat interactions with healthcare therapists, coaches, or other providers via the app?		X	
c. Does the app facilitate video chat interactions with healthcare therapists, coaches, or other providers via the app?		X	
d. Does the app facilitate teletherapy services via the app?		X	
e. Does the app facilitate group therapy services via the app?		X	
f. Does the app provide live support to a coach or counselor via the app?		X	
g. Does the app provide concierge mental health services via the app? Concierge services are personalized services patients can choose based on what they think works best for them. Some apps may provide acute concierge services to include additional help during times of crisis or high stress.		X	

3. Does the app provide a direct connection to 988 or other hotlines?

- Yes, and the connection works
- Yes, but the connection does not work
- No
- Unable to assess

4. Are the following functionalities supported by the app?

The following questions ask you to rate the functionality as either “comprehensive” or “not comprehensive.”

Comprehensive: Broad range of functionality that meets or exceeds the needs of the user and allows for flexibility

Not comprehensive: Rudimentary functionality that may not meet the needs of the user or provides limited choices and lacks flexibility

	Comprehensive	Not Comprehensive	Not Applicable
Mindfulness: Mindfulness is a therapeutic technique that includes elements of relaxation, breathing, and body exercise. It also includes techniques, such as meditation, guided positive imagery, grounding exercises, or progressive muscle relaxation.		X	
Journaling: Journaling can help users manage anxiety, reduce stress, and cope with mental health challenges. It can provide users with a way of identifying negative thoughts and behaviors and highlight positive aspects in their lives.		X	
Psychoeducation: Psychoeducation may be defined as the education of a person with a psychiatric disorder regarding the symptoms, treatments, and prognosis of that illness. Psychoeducation may also be targeted to the caregiver, family member, or loved one of the patient. Psychoeducation should be a brief personal intervention by a healthcare provider upon first diagnosis; however, sometimes, only written materials or online resources are provided to the patient. In the worst case scenario, patients receive no Psychoeducation from their provider. (Source: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD010823.pub2/full)		X	
Skill building: Skill building includes techniques for recognition of signs and symptoms of a problem, self-advocacy, stress management, emotional regulation, relapse prevention, promotion of sleep hygiene, etc. The app may include tips and advice on dealing with negative emotions and with behavior change. Skill building may require repetition, positive reinforcement, modeling, and practice.		X	
Screening: Historically, mental health care has relied on structured patient interviews and self-reported questionnaires for diagnosis. These can be used for self-evaluation, reporting, or to decide whether a patient should engage with a mental health professional. Please indicate if industry-validated screening questionnaires (e.g., GATT, PHQ-4, PHQ-9) are provided by the app. The standardized screening questionnaires used will generally be mentioned on the app developer's website.			X
Safety planning: Safety planning is designed to help individuals respond to escalation of suicidal thoughts and feelings, giving them a tailored list of coping behaviors and social support to use until the risk passes.			X

	Comprehensive	Not Comprehensive	Not Applicable
Sleep Hygiene: This includes features for promoting sleep hygiene, such as sleep diaries, lifestyle tracking, alarms, data synchronization with a wearable device (e.g., FitBit, Apple Watch), etc.			X
Automated Chatbots: Chatbots include conversational agents that can provide virtual therapy and/or social support.			X
Family/Caregiver Support: Caregivers take care of the day-to-day needs of patients. Apps can help caregivers monitor the mental state of the patient, identify the early signs of illness, track relapse and deterioration, and help the patient access services. The family/caregiver may also supervise treatment and provide emotional support to the patient.			X
Social and peer group interaction: Individuals with mental health conditions may find it hard to engage socially and may have a small social network/peer group. Apps may provide peer group interactions and/or community support discussion boards for such individuals.			X
Gamification: Gamification is the application of typical elements of game playing (e.g., point scoring, competition with others, rules of play) to other areas of activity. Gamification is used to encourage and reward positive changes.		X	
Personalization: Personalized treatment entails the selection of a therapy or treatment protocol based on a patient's profile, which may increase the likelihood of a successful outcome. Personalization can be used to tailor interactions initially, when the user starts to use the app, and also on a more regular basis. It is usually based on user information collected by the app, such as what activities produce positive emotions (e.g., physical exercise, talking to a friend).		X	
Self-Monitoring: Self-monitoring apps allow users to engage in symptom monitoring, mood tracking, and progress tracking through a treatment program, such as one for substance abuse, depression and anxiety, etc.	X		
Medication Adherence: Does the app support patients in their use of medication(s)? This could be through alerts, notifications, or other means.			X
Medication Delivery: Does the app support physical delivery of medications to patient(s)?			X

5. Please add any features supported by the app that are not listed above.

N/A

Post Administrative Questionnaire

This section contains some details about the app and evidence being reviewed.

1. Include all links to references for evidence (citations on PubMed, systematic reviews, websites, etc.) here.

None

2. Based on your review of the app, do you think it would have been useful to have some training or a tutorial about how to use it?

This is your subjective assessment based on your usage of the app.

- Yes, there was training available, and I needed it to use the app
- Yes, there was training available, but I didn't need it to use the app
- No, there was no training available, but I needed it to use the app
- No, there was no training available, and I didn't need it to use the app
- Don't know

3. Does the app have any help-related documentation available in the app itself (e.g., tooltips, general help)?

Base your answer on what was available on the app itself.

In-app help usually comes in three forms: 1) a separate page or pages of help within the app, with instructions; 2) popups that provide contextual help, displaying instructions relevant to the specific task that the user is attempting; 3) descriptions of app features of interest to the user. If at least two of these features is present, the help-related documentation would be considered to be comprehensive.

- Yes, there was comprehensive help available on the app
- Yes, there was some help available on the app
- The app help button took me to the website
- No help was available on the app

4. Does the app have a dedicated website that provides information about the app?

This may be part of a company website or a standalone website focused solely on the app.

- Yes
- No
- Unable to assess

5. Does the app have any help-related documentation available on its website?

- Yes, there was comprehensive help available
- Yes, there was some help available
- No, help was not available on the app

6. Please provide a subjective evaluation of the app in your own words.

- Do you think the assessment conducted here matches your subjective assessment? If not, please provide additional details.
- Do you think the risk posed by the app is captured accurately here?
- Do you think the app is technically sound?
- Do you think the therapeutic content provided by the app is rudimentary or substantive?
- Can this app help those that may be experiencing mental or behavioral health challenges?

This app appears to meet the assessment outlined here as well. Low risk but also no warning/disclaimers and connections to help. Technically sound, visually appealing and delightful app. Definetly feels created by app developers rather than mental health professionals.

7. Do you think the app could cause harm or have a negative impact on the user? Please elaborate.

No

8. Do you think the subjective risk assessment matches the risk assessment calculated in Section 1? Please elaborate.

Yes

9. Any additional comments?

10. Please note the time taken to complete your assessment using this framework.

50 minutes

Assessment Example, App C

Administrative Questionnaire

This section contains introductory questions about the app to be reviewed. The questions are mostly factual, rather than evaluative. Some of these questions may make it into the final Assessment Framework, based on the information we gather from the pilot.

Q 1. Reviewer Name

Reviewer 1

Q 2. Date of Evaluation

11/15/21

Q 3. App Name

App C

Q 4. App/ Company Website

This should only be a specific website developed by the app developer and not their pages on social media sites, such as Instagram or Facebook.

n/a

Q 5. Country of Origin

This information can be found in the Terms & Conditions on the website or in the app.

United Kingdom

Q 6. App Version Number

This information can be found on the app and from the App Store.

2.5.4

Q 7. Platform used for assessment by the reviewer

IOS

Q 8. OS Version Number on Evaluating Device

*For Android go to: Settings > About Phone > Software Information;
For iOS go to: Settings > General > Software Version.*

14.8.1

Q 9. Describe the goal of the app in your own words. Do not use any of the marketing phrases/lingo used by the app developer.

This should be based on the app description on the website.

to track moods, learn about mental health, take mental health tests, track mood, stay motivated and relaxed

Q 10. Has the app been approved by any regulating authority, such as FDA?

This information can be found on the app developer website.

no

Q 11. Does the app have multiple revenue models, such as freemium, in-app purchases, etc.?

- Yes
- No

Q 12. If the app has multiple versions (e.g., free, freemium, paid), which version did you evaluate? If the app had only one revenue model, please choose “not applicable” here.

- Free
- Paid version
- Paid by insurance agency, employer, or healthcare provider
- Not applicable
- Unable to assess

Q 13. Does the app require prior authorization from a healthcare institution/insurance provider/college/employer or other institution for access?

- Yes
- No
- Unable to assess

Q 14. If you were unable to assess the app, please specify the reason.

n/a

Note: Questions about privacy, security, and informed consent are covered in both Section 1 and Section 2 of the framework. Please make a note of all the disclaimers, warnings, and privacy and security statements and agreements when you first log into the app, so you are able to answer the questions that come later in the framework.

Section 1: Risks and Mitigation Strategies

Section 1 aims to evaluate the potential risk posed by the app, and evidence of efficacy, and safety features. The burden of evidence to prove efficacy and safety is dependent on the goals of the app, the target audience, and the severity of the mental health condition.

The categories of questions in this section are App Integrity, Risk Assessment, Evidence, Linkage to Care, and Access to Crisis Services. If apps do not meet the threshold for safety and credibility, that is flagged in the assessment.

A. App Integrity

The information requested in these questions can be found on the App Store, the app's website, or via Google Search.

1. Was the version of the app you are reviewing updated in the last 6 months?

- Yes
- No
- Unable to assess

2. Does the app provide a privacy and security agreement to be reviewed and agreed to by the user?

The agreement should be accessible through a link on the app or on the mobile app. Reviewers should not look for this information in other locations, such as the developer's website.

- Yes, there is a privacy and security agreement to be agreed to by the user
- No, there is not a privacy and security agreement
- Unable to assess or unable to access

3. Does the app provide warnings and disclaimers (e.g., the limitations of the app related to medical liability)?

Reviewers should look for this information on the mobile app ONLY.

- Yes, warnings and/or disclaimers are provided
- No, warnings and/or disclaimers are not provided
- Unable to assess
- Not applicable

4. Has the app been endorsed by, or is it being used by, a government agency or trusted mental health professional association?

Endorsement and/or usage helps us establish the credibility of the app.

Endorsement means to give one's approval to, especially officially by an organization. An app merely being reviewed by the American Psychological Association (APA), or a similar organization, does not mean it has been approved by them. Endorsement should be for the app and not the company developing the app or the domain (e.g., teletherapy). Self-reported endorsements are acceptable.

Usage by an institution signifies that an app is made available to all its members. For example, Veterans Affairs uses iBlueButton to deliver critical health care information to veterans.

Examples of government agencies include federal, state and local agencies such as CMS, Veterans Affairs, Virginia Department of Health. Examples of trusted mental health professional associations include Substance Abuse and Mental Health Services Administration (SAMHSA), American Psychological Association (APA), and National Alliance on Mental Illness (NAMI).

- Yes, the app has been endorsed by one or more mental health associations, government agencies, or non-government bodies (examples may include but are not limited to government agencies, such as Centers for Medicare & Medicaid Services (CMS) or Veterans Affairs (VA); non-government bodies, such as APA; and insurance or healthcare institutions)
- No, the app has not been endorsed
- Unable to assess, please specify reason: _____

App Integrity Assessment

Question	Response	App Integrity Level
Does the app satisfy the Integrity Assessment?	<input checked="" type="radio"/> Responses to questions 1-3 are all Yes OR response to question 4 is Yes	App Integrity is "High"
Please pick only 1 response.	<input type="radio"/> Response to any of the questions 1-3 is No AND response to 4 is No	App Integrity is "Low"

B. Risk Assessment

This category aims to determine the level of risk posed by the app. For example, if the app is used by someone with a clinically diagnosed condition, such as Schizophrenia, then the risk posed by the app may be higher than if the person has mild anxiety. Similarly, if the app targets children and adolescents, then the potential risk associated with the app may be higher than if it targeted adults with no impairment or mental health diagnosis.

*For the purposes of this report, either minors (younger than 18 years of age) or adults with a moderate or severe level of functional impairment that causes them to need support from a caregiver are considered to be part of a **vulnerable population**.*

The information requested in these questions can be found within the app/app website.

5. Is the app intended for use by adults who may have an illness or disability that impacts their decisionmaking ability? Severe autism or severe dementia are examples of illnesses or disabilities that may affect decisionmaking ability.

For the purposes of this assessment, the population that may have an illness or disability that impacts their decision making ability is included in the "vulnerable population" group. Adults can usually make medical decisions, such as consenting to treatment, on their own. In some cases, adults need someone else to make medical decisions for them if their decisionmaking ability is affected by an illness or disability. This is sometimes referred to as "substituted consent." A higher level of evidence/protection is needed for the use of specific apps under these circumstances.

- Yes
- No
- Unable to assess

6. Does the app claim that it is intended for use by minors (i.e., people younger than 18 years of age)?

For the purposes of this assessment, this population is included in the "vulnerable population" group. The app is "intended for use" by a population if it is clearly specified on the app or website that this population is its target user group. Do not look at the App Store Age Ratings or Terms and Conditions for this information.

- Yes, for use by minors
- Yes, for use by both minors and adults
- No, not for use by minors
- Unable to assess

7. Does the app claim to provide standalone treatment for any mental health condition?

A standalone treatment might be some form of psychotherapy or other any other medical care provided by the app independently, rather than in conjunction with a licensed healthcare practitioner.

Functionality, such as mindfulness training, skills training, symptom or mood tracking, and social support, would not be considered standalone treatment.

- Yes, the app provides standalone treatment

- No, the app does not provide standalone treatment
- Unable to assess

C. Evidence

It is important that the app has a solid clinical foundation. If the app targets a vulnerable population, the burden of evidence to prove efficacy and safety is greater.

The information requested in these questions can be found on the app website (look at customer testimonials or resources/blog for anecdotal information) and on PubMed (search PubMed as outlined in the training document). Make sure to save all the references in the app review document.

8. Has the app been evaluated for efficacy/effectiveness through a scientifically validated study?

This can be determined by evaluating whether a research study about the app has been published in a peer-reviewed journal. The research should be specific to the mental health component of the app.

- Yes, there is strong research support for the app (i.e., at least one published paper in a peer-reviewed journal that uses a randomized trial that shows efficacy or effectiveness)
- Yes, there is some research support for the app (i.e., at least one published paper in a peer-reviewed journal that uses single-case design, quasi-experimental methods demonstrating efficacy)
- No, there is no research support for the app

9. Is the app based on or does it use an evidence-based strategy, such as Cognitive Behavioral Therapy (CBT), Dialectical Behavior Therapy (DBT), or evidence-based guidelines?

Evidence-based medicine aims to assess the strength of proof behind medical interventions in terms of risks and benefits and, therefore, can be used to inform clinical decisionmaking on both an individual and a population basis.

- Yes, the app reports using an evidence-based strategy/guidelines to achieve its goals
- No, the app does not report the use of evidence-based strategy/guidelines to achieve its goals (or there are no goals described)
- Unable to assess

D. Linkage to Care

This category evaluates the linkages to healthcare providers, caregivers. This information can be found within the app/app website.

10. Does the app facilitate remote monitoring of the patient or send alerts to a clinician/clinical care team or caregiver?

This question seeks to assess whether a healthcare provider or caregiver can monitor the user's health data in real time AND/OR receive alerts triggered by certain pre-specified high-risk events. Functionality to export data into a file or an email does not count as facilitation of remote monitoring.

- Yes
- No
- Unable to assess

10.a. If YES, please specify who can monitor the health of the patient, either through alerts or by other means?

- Healthcare provider only
- Caregiver only
- Both healthcare provider and caregiver
- Other
- Not applicable

E. Access to Crisis Services

This section evaluates whether the app provides access to emergency sources of information.

This information can be found within app/app website. There should be a caregiver or healthcare provider interface within the app. Exporting data into a file or an email would not be considered facilitation of remote monitoring.

11. Does the app provide users with information about resources that can be reached in case of an emergency? This information should be available on the mobile app. Reviewers should not look for this information on the website.

- Yes (e.g., a crisis hotline, 911, nearest emergency room services.)
- No, the app provides no information on emergency services
- Unable to assess

12. Evidence and Linkage to Care: If the app can be used by individuals that require substituted consent OR by minors, then is consent sought from either a caregiver/parent/legal guardian?

This should be assessed at the time of creating a login for the app.

- Yes
- No
- Unable to assess
- Not applicable

Risk Assessment

What is the Risk Category for the App?

A lower number signifies a lower risk level.

Risk Level 1: Apps rated as Risk Level 1 do not target a vulnerable population AND do not provide standalone treatment. These are likely to be apps that provide functions such as mindfulness, meditation, and wellness resources. No specific safety checks are needed for these apps.

- Risk Level is 3:** Response to EITHER Question 5 or Question 6 is Yes AND response to Question 7 is Yes, *go to Safety Assessment Step 3*
- Risk Level is 2:** Response to EITHER Question 5 or Question 6 is Yes OR response to Question 7 is Yes, *go to Safety Assessment Step 3*
- Risk Level is 1:** None of the above, *go to Safety Assessment Step 3*

Safety Assessment

Step 1: Preliminary Safety Assessment for Risk Levels 2 & 3

Risk Level 2	<input type="radio"/> [If responses to Questions 9-11 are Yes AND response to Question 10 is Yes or No, supporting evidence-base available]	PRELIM SECURITY CHECK PASSED , <i>go to Safety Assessment Step 2</i>
	Else	PRELIM SECURITY CHECK FAILED , <i>go to Safety Assessment Step 2</i>
Risk Level 3	<input type="radio"/> [If responses to Questions 8-11 are Yes, supporting evidence-base available]	PRELIM SECURITY CHECK PASSED , <i>go to Safety Assessment Step 2</i>
	Else	PRELIM SECURITY CHECK FAILED , <i>go to Safety Assessment Step 2</i>

Step 2: Safety Assessment for Vulnerable Population

If response to Question 5 and/or Question 6 is Yes,

- If response to Question 12 is **Yes**, then **SAFE for vulnerable populations**, go to *Safety Assessment Step 3*
- If response to Question 12 is **No**, then **NOT SAFE for vulnerable populations**, go to *Safety Assessment Step 3*

else go to *Step 3 under Safety Assessment*

STEP 3: Safety Assessment

RISK LEVEL	PRELIM SAFETY CHECK	VULNERABLE POPULATION	APP SAFETY
Risk Level 1	N/A	N/A or SAFE for Vulnerable Populations	Passed Safety Check
	N/A	Not Safe for Vulnerable Populations	Failed Safety Check
Risk Level 2	PASS	N/A or SAFE for Vulnerable Populations	Passed Safety Check
		Not Safe for Vulnerable Populations	Failed Safety Check
	FAIL	--	Failed Safety Check
Risk Level 3	PASS	N/A or SAFE for Vulnerable Populations	Passed Safety Check
		Not Safe for Vulnerable Populations	Failed Safety Check
	FAIL	--	Failed Safety Check

Section 2: Function

This section is focused on descriptive aspects, such as accessibility, usability, privacy/security, AI, cultural competence, and informed consent.

A. Accessibility Features

Accessibility features are meant to make the app easier to navigate, especially for those with disabilities. Common accessibility features include text-to-speech, closed-captioning, and keyboard shortcuts.

1. Which of the phone's accessibility features work within the app?

Reviewers will have to activate accessibility features on their phone under Settings in order to test this.

Select all that apply:

- Text adjustment feature
- Colorblind color scheme feature
- Text-to-speech feature
- None of the phone's accessibility features work in the app
- Unable to assess, please provide a reason: _____

2. Are there additional accessibility features that are provided by the app?

These are accessibility features that are provided within the app but cannot be set as an option under Settings.

Select all that apply:

- Adaptation of audio/video content with transcriptions or captions
- Tapping and other gestures are configurable
- Contrast text coloring in the content
- Screen reader
- No additional accessibility features provided by the app
- Other features not provided here (please specify): _____
- Unable to assess, please provide a reason: _____

B. App Information

General info about the app and user ratings.

The information requested in these questions can be obtained by going to the appropriate app store using your browser.

3. Does the app work on Apple(iOS)?

- Yes
- No

4. What are the number of reviews on the iOS App Store?

The iOS App Store can be accessed through a browser on your phone, tablet, or computer.

7.2k

5. What is the app rating (number of stars) on Apple Store?

This provides users' perspectives of the app.

4.6

6. Does the app work on Android?

Yes

No

7. What are the number of reviews on Google Play Store.

61

8. What is the app rating (number of stars) on Google Play Store?

The Google Play Store can be accessed through a browser on your phone, tablet, or computer.

3.6

C. Costs

App costs could be upfront, in the form of a monthly/annual subscription, or through freemium services that require in-app purchasing. Some apps require payment from the user, while others are reimbursed by the healthcare provider or insurance.

The information requested in these questions may be found in the app, in the App Store, and on the app site.

9. What is the business model for the app? More than one option may be applicable here. An app that provides a free trial for a limited duration is not considered to be free.

If the app provides a payment gateway to pay a therapist or healthcare provider, those costs should not be included.

- Free (no upfront fee for the app; additional packages/services may be offered for a fee)
- Upfront fee (a onetime cost for accessing the app; additional packages/services may be offered for additional fees)
- In-app purchases (additional packages/services available for a fee in addition to or in lieu of an upfront cost, e.g., concierge services)
- Subscription (payment for services on a monthly/quarterly/annual basis)
- Reimbursed by healthcare providers/insurers/employers
- Other, please specify: _____

10. Does the app provide a free or freemium model? An app that provides a free trial for a limited duration is not considered to be free.

- Free
- Freemium
- No free or freemium version

11. What is the estimated annual cost of the app for the paid version?

If the app provides a free or freemium model, then select “Not applicable” in the list below.

- Under \$50
- \$50-250
- Over \$250
- Not applicable

12. If the app includes paid service(s), does it provide CPT code(s) for insurance reimbursement?

Current Procedural Terminology (CPT) are unique 5-digit codes for medical services or procedures assigned by the American Medical Association which are used throughout the U.S. medical system. Insurance companies decide on which CPT codes can be used for insurance reimbursement.

- Yes
- No or the app has no paid services
- Unable to Assess, please provide a reason: _____

D. Organizational Credibility

Organizational credibility is meant to determine whether the app comes from a trusted source.

13. Who is the developer of the app?

This information can be found on the app website or in the Terms of Service on the app. More than one response may apply. For example, two or more institutions may be involved in the development of the app; or, for example, if an app was developed by Johns Hopkins University, then “Non-profit institution” and “Academic institution” would apply.

- Government
- For-profit company
- Non-profit institution
- Healthcare institution
- Academic institution
- Insurance company
- Independent developer/s
- Unable to assess
- Other, please specify: _____

14. Does the app have any consumer bureau complaints or lawsuits pending?

This information can be found here: Better Business Bureau: <https://www.bbb.org/us/ca/san-francisco/profile/mobile-apps/calmcom-inc-1116-877519/complaints> ; Google Search: <Consumer Complaints> <app name> <app/mobile app>; Google Search: <law suits> <app> <app/mobile app>

- Yes
- No
- Unable to assess

E. Evidence & Clinical Foundation

The information requested in these questions can be found on the app website.

15. Does the app appear to do what it claims to do?

Look at the website to see the claims made by the app and compare them to what you see in the app.

- Yes, the app provides the functionality it claims on its website
- The app provides some of the functionality it claims on its website
- No, the app does not provide the functionality it claims on its website
- Unable to assess

F. Privacy & Security

Security is about the safeguarding of data and unauthorized access of the app. Privacy is about the safeguarding of the user's identity.

The information requested in these questions should be available in the app in the Terms & Conditions or Privacy & Security agreement. Do not look at information provided on the website.

16. Does the app claim it meets HIPAA [or analogous national standard for protected health information (PHI)]?

- Yes
- No

17. Does the app claim it meets COPPA [or analogous national standard for protected health information (PHI) for minors younger than 13 years of age]?

The Children's Online Privacy Protection Act (COPPA) applies to the online collection of personal information by persons or entities under U.S. jurisdiction about children younger than 13 years of age.

- Yes
- No

18. Does the app report sharing or selling of data for research or commercial purposes?

Information used internally within the app itself to make better recommendations to the user does not count as sharing or selling. Examples of apps that may be sharing or selling data include those that provide evidence based drug and dosage information, clinical decision support systems where you share conditions, symptoms and medication information, medication adherence apps, etc.

- Yes
- No
- Unable to assess

19. If your response to question 18 is “YES”, please answer the following question:

If the app reports sharing or selling for research or commercial purposes, are the data de-identified?

This question is targeted at understanding if the data is de-identified or not. Information used internally, within the app itself, to make better recommendations to the user does not count as sharing or selling.

- Yes, the app reports that data is de-identified
- No, the app does not report data is de-identified
- Unable to assess

20. If the app has the capability to read/write to an electronic health record management system (EHRs) or other healthcare systems, does it use industry standards for secure interoperability (e.g., FHIR, SMART, OAuth 2.0, TLS 1.2)?

This information may need to be deduced. If the app has existing integrations with known EHRs such as EPIC, it can be concluded that it follows industry standards. This information may also need to be identified through a review of the app developer webpage.

- Yes, the app uses industry standards for interoperability
- No, the app does not use industry standards for interoperability
- Not applicable, because it does not read/write to EHRs
- Unable to assess

G. Informed Consent

Informed consent is permission granted by the user after information about the potential use or disclosure of the information collected by the app is provided. Most apps tend to have a disclosure list that is long and hard to understand. There are best practices for ensuring that users understand exactly what they are agreeing to before they “Agree” to the app’s privacy and security practices. The following questions evaluate whether the app follows these best practices.

The information requested in these questions can be found within the app.

21. Please assess the level of informed consent enabled by the app.

Providing consent is often one of the first things users are asked to do after downloading an app. Traditional written consent forms (e.g., Terms of Service, Privacy and Security agreement) are often too long, difficult to understand, and overly focused on legal concerns. It is important that apps enable informed consent for prospective users in a manner that is easy to understand.

- No informed consent
- Does not simplify informed consent
- Average: Focuses on the essential by a) providing a narrative focused on the most salient information, b) limiting concepts to one per screen, and c) following national plain language and health literacy guidance
- Good: Focuses on the essential and organizes content deliberately by a) prioritizing key words/concepts presented on each screen, b) providing information tiers for conceptual elaboration, and c) enabling participants to navigate to their desired level of detail
- Excellent: Focuses on the essential, organizes content deliberately, and encourages engagement through interactive elements

22. What is the format of the data privacy and security consent process followed in the app?

*An opt-in consent process requires **explicit consent from the user** before the collection and processing of their personal data. It refers to a **positive action** taken by the user indicating that they agree to the use of their personal data. An opt-out consent process **does not require the user’s consent** prior to the collection and processing of their personal data. It refers to the process by which users **withdraw their consent** to the use of their personal data.*

- Opt-in for data to be shared, the default choice is opt-out
- Opt-out of data sharing, the default choice is opt-in
- No choice to opt-out of data sharing, without explicit consent you can’t use the app
- Unable to assess

H. Cultural Competence

Cultural competence is defined as the ability to understand, appreciate, and account for different cultures or belief systems based on race, ethnicity, income strata, religious beliefs, etc. It is important to assess if the app captures these differences and whether it provides personalized care that takes these cultural differences into account.

The information requested in these questions can be found within the app or on the app website.

23. Does the app report developing and testing the app for specific cultural group/s?

According to the U.S. Department of Health and Human Services (HHS), "cultural competence is the integration and transformation of knowledge about individuals and groups of people into specific standards, policies, practices, and attitudes used in appropriate cultural settings to increase the quality of services; thereby producing better outcomes." For this assessment, groups with lived experiences such as pregnant teens and survivors of gender-based violence would be considered to be a specific cultural group.

- Yes
- No
- Unable to assess
- Other, please specify: _____

24. If the app reports developing and testing the app for specific cultural groups, is there published documentation (e.g., website, papers) of the process taken to incorporate information that is specific to specified cultures?

Select all that apply:

- Website
- Documentation (on the app or website)
- Published scientific papers
- Unable to assess
- Other, please specify: _____
- Not applicable

If the app integrates culturally specific groups, please name the groups.

25. Is gender-inclusive language employed? When asking the user’s gender, is there an option to self-describe, an option to decline to answer, use of scientifically correct terms for gender (e.g., man, woman, nonbinary)?

This information can be found while assessing the app and is generally encountered when a user creates an account on the app.

Select all that apply:

- An option for use of personal pronouns (e.g., they/them, he/his, she/her)
- An option to specify another gender
- An option to decline to answer
- Use of scientifically correct terms for gender, such as man, woman, non-binary
- Unable to assess
- Not applicable

26. If the app was tested in a study, what was the percentage of non-white participants?

For clinical trials, this information may be found on the clinicaltrials.gov page of the trial under Study Results (include “Other” as non-white). If the study was a usability or other type of study, this information may be found on the developer website.

- Less than 30%
- Between 30 – 50%
- More than 50%
- No disaggregate data available
- No information available

I. Usability

Usability can be described as the capacity of a system to provide a condition for its users to perform the tasks safely, effectively, and efficiently. It is important that the experience be engaging and pleasing, otherwise users are likely to stop using the app.

The information requested in these questions can be found within the app.

27. Does the app work offline?

Choose Yes if parts of the app or the entire app work offline.

- Yes
- No
- Unable to assess

28. Performance: How accurately/fast do the app features (functions) and components (buttons/menus) work?

- App is broken; no/insufficient/inaccurate response (e.g., crashes/bugs/broken features)
- Some functions work, but app lags or contains major technical problems
- App works overall; some technical problems need fixing or app is slow at times
- Mostly functional with minor/negligible problems
- Perfect/timely response; no technical bugs found or contains a ‘loading time left’ indicator, if relevant

29. What languages are supported by the application?

This information can be found in the App Store.

Select all that apply:

- English
- Spanish
- French
- German
- Other, please specify: _____

30. Customization: Does the app allow customization of settings and preferences by the user (e.g., sound, content, notifications)?

- Does not allow any customization or requires setting to be input every time
- Allows little customization and that limits the app’s functions
- Basic customization to function adequately
- Allows numerous options for customization
- Allows complete tailoring to the user’s characteristics/preferences, remembers all settings

31. Target group: Is the app content (visuals, language, design) appropriate for the target audience?

- Completely inappropriate, unclear, or confusing
- Mostly inappropriate, unclear, or confusing
- Satisfactory, but not specifically designed for the target audience, may be inappropriate/unclear/confusing at times
- Designed for the target audience, with minor issues
- Designed specifically for the target audience; no issues found

32. Layout: Is the arrangement and size of buttons, icons, menus, and content on the screen appropriate?

- Very bad design: cluttered, some options impossible to select/locate/see/read
- Bad design: random, unclear, some options difficult to select/locate/see/read
- Satisfactory: few problems with selecting/locating/seeing/reading items
- Mostly clear: able to select/locate/see/read items
- Professional: simple, clear, orderly, logically organized

33. Graphics: How high is the quality/resolution of graphics used for buttons, icons, menus, and content?

- Graphics appear amateur, very poor visual design: disproportionate, stylistically inconsistent
- Low quality/low resolution graphics, low quality visual design: disproportionate
- Satisfactory quality graphics and visual design: generally consistent in style
- High quality/resolution graphics and visual design: mostly proportionate, consistent in style
- Very high quality/resolution graphics and visual design: proportionate, consistent in style throughout

34. Visual appeal: How good does the app look?

- Ugly/unpleasant to look at: poorly designed, clashing, mismatched colors
- Bad: poorly designed, bad use of color, visually boring
- Satisfactory: average, neither pleasant nor unpleasant
- Pleasant: seamless graphics, consistent and professionally designed
- Beautiful: very attractive, memorable, stands out; use of color enhances app features/menus

35. Does the app have advertising?

- Yes
- No

36. If app has advertising, is the advertising intrusive and distracting?

- Advertising is neither intrusive nor distracting
- The advertising is both intrusive and distracting
- Not applicable

37. Ease of use: How easy is it to learn how to use the app; how clear are the menu labels, icons, and instructions?

- No/limited instructions, menu labels and icons are confusing, complicated
- Takes a lot of time or effort
- Takes some time or effort
- Easy to learn (or has clear instructions)
- Able to use app immediately, intuitive and simple (no instructions needed)

38. Navigation: Does moving between screens make sense? Does the app have all necessary links between screens?

- No logical connection between screens at all/navigation is difficult
- Understandable after a lot of time/effort
- Understandable after some time/effort
- Easy to understand/navigate
- Perfectly logical, easy, clear and intuitive screen flow throughout, and/or has shortcuts

39. Gestural design: Do taps/swipes/pinches/scrolls make sense? Are they consistent across all components/screens?

- Completely inconsistent/confusing
- Often inconsistent/confusing
- Satisfactory with some inconsistencies/confusing elements
- Mostly consistent/intuitive with negligible problems
- Perfectly consistent and intuitive

40. Content: Is the app copy well written and relevant to the goal and/or topic of the app?

- There is no information within the app
- Irrelevant/inappropriate/incoherent/incorrect
- Barely relevant/appropriate/coherent, may be incorrect
- Satisfactory with respect to relevance/appropriateness/coherence and appears to be correct
- Relevant/appropriate/coherent/correct
- Highly relevant/appropriate/coherent/correct
- Not applicable

41. Is there any evidence to show the average duration of use of the app by users?

This information may be available on the app website or in a paper published in a scientific journal. Do not look at user reviews on the app website, in the App Store, or in news articles.

- On average, users have engaged with the app for 30 days or less
- On average, users have engaged with the app for more than a month but less than 6 months
- On average, users have engaged with the app for more than 180 days
- No evidence available

J. Remote Monitoring

Remote patient monitoring enables monitoring of patients outside of conventional clinical settings through real-time access to patient health data. The patient's provider may have a schedule to monitor patient data or may receive alerts about their patient's health.

The information requested in these questions can be found within the app/app website.

42. If remote monitoring is feasible, how does the provider access the data?

- Provider has access to patient information through the app
- Provider has access to the patient information through the EMR
- Provider has no access to patient information
- Not applicable
- Unable to assess
- Other, please specify: _____

43. Does the app provide alerts to the provider to notify them of a clinical event that may require action on their part?

Two-way messaging capability with a healthcare provider or concierge may be considered to be an alerting mechanism.

- Yes
- No
- Unable to assess

44 . Can the app share data with wearables like Apple Watch and Fitbit?

This option will either pop up on the app during sign-up or be mentioned on the app developer's website.

- Yes
- No
- Unable to assess

K. Access to Crisis Response Services

This section evaluates whether the app provides access to emergency sources of information.

The information requested in this question can be found within the app or on the app website.

45. Does the app connect the user automatically to resources in case of a crisis situation or emergency?

- Yes
- No
- Unable to assess

L. Artificial Intelligence (AI)

Apps often use artificial intelligence (AI) for a variety of tasks including risk-prediction, differential diagnoses, and personalization of health content. The following questions are important because we want to gauge the potential of the app to cause harm and also to determine whether the app's algorithms are being updated based on user input.

The information requested in these questions can be found in the app, on the app website, or from evidence found through a literature search.

46. Does the app use AI? If yes, please respond to questions 47 and 48.

- Yes
- No

47. How is AI being used in the app?

For example, AI may be used in the app to automate simple tasks, such as scheduling, or it may be used for more complex functionality, such as to personalize chatbots.

48. Is there any evidence to suggest that the app uses data from user interactions to improve precision on AI models?

For example, after an interaction, the app may ask the user for feedback on the automated response from the app to see if it was appropriate. If the response was not correct, the app may ask for further feedback from the user so the algorithms can be refined.

- Yes, there is evidence to suggest that the AI models are updated based on feedback
- No, there is no evidence to suggest that the AI models are updated based on feedback
- Unable to assess

Section 3: Mental Health App Features

This section attempts to capture the features specific to mental health present in the app. This is not meant to be an exhaustive list. Only some of these features may apply to the app being reviewed.

The information requested in these questions can be found in the app, on the app website or in published literature.

1. List mental health symptom(s) and/or condition(s) addressed by the app.

depression, anxiety, stress

2. Please answer the following questions about the app features.

Question	Yes	No	Unable to assess
a. Does the app facilitate text messaging interactions with healthcare therapists, coaches, or other providers via the app?		x	
b. Does the app facilitate audio chat interactions with healthcare therapists, coaches, or other providers via the app?		x	
c. Does the app facilitate video chat interactions with healthcare therapists, coaches, or other providers via the app?		x	
d. Does the app facilitate teletherapy services via the app?		x	
e. Does the app facilitate group therapy services via the app?		x	
f. Does the app provide live support to a coach or counselor via the app?		x	
g. Does the app provide concierge mental health services via the app? Concierge services are personalized services patients can choose based on what they think works best for them. Some apps may provide acute concierge services to include additional help during times of crisis or high stress.		x	

3. Does the app provide a direct connection to 988 or other hotlines?

- Yes, and the connection works
- Yes, but the connection does not work
- No
- Unable to assess

4. Are the following functionalities supported by the app?

The following questions ask you to rate the functionality as either “comprehensive” or “not comprehensive.”

Comprehensive: Broad range of functionality that meets or exceeds the needs of the user and allows for flexibility

Not comprehensive: Rudimentary functionality that may not meet the needs of the user or provides limited choices and lacks flexibility

	Comprehensive	Not Comprehensive	Not Applicable
Mindfulness: Mindfulness is a therapeutic technique that includes elements of relaxation, breathing, and body exercise. It also includes techniques, such as meditation, guided positive imagery, grounding exercises, or progressive muscle relaxation.		x	
Journaling: Journaling can help users manage anxiety, reduce stress, and cope with mental health challenges. It can provide users with a way of identifying negative thoughts and behaviors and highlight positive aspects in their lives.			x
Psychoeducation: Psychoeducation may be defined as the education of a person with a psychiatric disorder regarding the symptoms, treatments, and prognosis of that illness. Psychoeducation may also be targeted to the caregiver, family member, or loved one of the patient. Psychoeducation should be a brief personal intervention by a healthcare provider upon first diagnosis; however, sometimes, only written materials or online resources are provided to the patient. In the worst case scenario, patients receive no Psychoeducation from their provider. (Source: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD010823.pub2/full)	x		
Skill building: Skill building includes techniques for recognition of signs and symptoms of a problem, self-advocacy, stress management, emotional regulation, relapse prevention, promotion of sleep hygiene, etc. The app may include tips and advice on dealing with negative emotions and with behavior change. Skill building may require repetition, positive reinforcement, modeling, and practice.			x
Screening: Historically, mental health care has relied on structured patient interviews and self-reported questionnaires for diagnosis. These can be used for self-evaluation, reporting, or to decide whether a patient should engage with a mental health professional. Please indicate if industry-validated screening questionnaires (e.g., GATT, PHQ-4, PHQ-9) are provided by the app. The standardized screening questionnaires used will generally be mentioned on the app developer's website.	x		
Safety planning: Safety planning is designed to help individuals respond to escalation of suicidal thoughts and feelings, giving them a tailored list of coping behaviors and social support to use until the risk passes.			x

	Comprehensive	Not Comprehensive	Not Applicable
Sleep Hygiene: This includes features for promoting sleep hygiene, such as sleep diaries, lifestyle tracking, alarms, data synchronization with a wearable device (e.g., FitBit, Apple Watch), etc.			x
Automated Chatbots: Chatbots include conversational agents that can provide virtual therapy and/or social support.			x
Family/Caregiver Support: Caregivers take care of the day-to-day needs of patients. Apps can help caregivers monitor the mental state of the patient, identify the early signs of illness, track relapse and deterioration, and help the patient access services. The family/caregiver may also supervise treatment and provide emotional support to the patient.			x
Social and peer group interaction: Individuals with mental health conditions may find it hard to engage socially and may have a small social network/peer group. Apps may provide peer group interactions and/or community support discussion boards for such individuals.			x
Gamification: Gamification is the application of typical elements of game playing (e.g., point scoring, competition with others, rules of play) to other areas of activity. Gamification is used to encourage and reward positive changes.			x
Personalization: Personalized treatment entails the selection of a therapy or treatment protocol based on a patient's profile, which may increase the likelihood of a successful outcome. Personalization can be used to tailor interactions initially, when the user starts to use the app, and also on a more regular basis. It is usually based on user information collected by the app, such as what activities produce positive emotions (e.g., physical exercise, talking to a friend).			x
Self-Monitoring: Self-monitoring apps allow users to engage in symptom monitoring, mood tracking, and progress tracking through a treatment program, such as one for substance abuse, depression and anxiety, etc.		x	
Medication Adherence: Does the app support patients in their use of medication(s)? This could be through alerts, notifications, or other means.			x
Medication Delivery: Does the app support physical delivery of medications to patient(s)?			x

5. Please add any features supported by the app that are not listed above.

n/a

Post Administrative Questionnaire

This section contains some details about the app and evidence being reviewed.

1. Include all links to references for evidence (citations on PubMed, systematic reviews, websites, etc.) here.

2. Based on your review of the app, do you think it would have been useful to have some training or a tutorial about how to use it?

This is your subjective assessment based on your usage of the app.

- Yes, there was training available, and I needed it to use the app
- Yes, there was training available, but I didn't need it to use the app
- No, there was no training available, but I needed it to use the app
- No, there was no training available, and I didn't need it to use the app
- Don't know

3. Does the app have any help-related documentation available in the app itself (e.g., tooltips, general help)?

Base your answer on what was available on the app itself.

In-app help usually comes in three forms: 1) a separate page or pages of help within the app, with instructions; 2) popups that provide contextual help, displaying instructions relevant to the specific task that the user is attempting; 3) descriptions of app features of interest to the user. If at least two of these features is present, the help-related documentation would be considered to be comprehensive.

- Yes, there was comprehensive help available on the app
- Yes, there was some help available on the app
- The app help button took me to the website
- No help was available on the app

4. Does the app have a dedicated website that provides information about the app?

This may be part of a company website or a standalone website focused solely on the app.

- Yes
- No
- Unable to assess

5. Does the app have any help-related documentation available on its website?

- Yes, there was comprehensive help available
- Yes, there was some help available
- No, help was not available on the app

6. Please provide a subjective evaluation of the app in your own words.

- Do you think the assessment conducted here matches your subjective assessment? If not, please provide additional details.
- Do you think the risk posed by the app is captured accurately here?
- Do you think the app is technically sound?
- Do you think the therapeutic content provided by the app is rudimentary or substantive?
- Can this app help those that may be experiencing mental or behavioral health challenges?

Very cute app with comprehensive, basic functionality

7. Do you think the app could cause harm or have a negative impact on the user? Please elaborate.

no

8. Do you think the subjective risk assessment matches the risk assessment calculated in Section 1? Please elaborate.

yes

9. Any additional comments?

n/a

10. Please note the time taken to complete your assessment using this framework.

30 minutes