

BMJ Open Evaluating evaluation frameworks: a scoping review of frameworks for assessing health apps

Sarah Lagan, Lev Sandler, John Torous 

To cite: Lagan S, Sandler L, Torous J. Evaluating evaluation frameworks: a scoping review of frameworks for assessing health apps. *BMJ Open* 2021;**11**:e047001. doi:10.1136/bmjopen-2020-047001

► Prepublication history and additional material for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2020-047001>).

Received 17 November 2020

Revised 06 March 2021

Accepted 10 March 2021

ABSTRACT

Objectives Despite an estimated 300 000 mobile health apps on the market, there remains no consensus around helping patients and clinicians select safe and effective apps. In 2018, our team drew on existing evaluation frameworks to identify salient categories and create a new framework endorsed by the American Psychiatric Association (APA). We have since created a more expanded and operational framework Mhealth Index and Navigation Database (MIND) that aligns with the APA categories but includes objective and auditable questions (105). We sought to survey the existing space, conducting a review of all mobile health app evaluation frameworks published since 2018, and demonstrate the comprehensiveness of this new model by comparing it to existing and emerging frameworks.

Design We conducted a scoping review of mobile health app evaluation frameworks.

Data sources References were identified through searches of PubMed, EMBASE and PsychINFO with publication date between January 2018 and October 2020.

Eligibility criteria Papers were selected for inclusion if they meet the predetermined eligibility criteria—presenting an evaluation framework for mobile health apps with patient, clinician or end user-facing questions.

Data extraction and synthesis Two reviewers screened the literature separately and applied the inclusion criteria.

The data extracted from the papers included: author and dates of publication, source affiliation, country of origin, name of framework, study design, description of framework, intended audience/user and framework scoring system. We then compiled a collection of more than 1701 questions across 79 frameworks. We compared and grouped these questions using the MIND framework as a reference. We sought to identify the most common domains of evaluation while assessing the comprehensiveness and flexibility—as well as any potential gaps—of MIND.

Results New app evaluation frameworks continue to emerge and expand. Since our 2019 review of the app evaluation framework space, more frameworks include questions around privacy (43) and clinical foundation (57), reflecting an increased focus on issues of app security and evidence base. The majority of mapped frameworks overlapped with at least half of the MIND categories. The results of this search have informed a database (apps.digitalpsych.org) that users can access today.

Conclusion As the number of app evaluation frameworks continues to rise, it is becoming difficult for users to

Strengths and limitations of this study

- This scoping review is the largest and most up to date review and comparison of mobile health app evaluation frameworks.
- The analysis highlighted the flexibility and comprehensiveness of the Mhealth Index and Navigation Database (MIND) framework, which was used as a reference framework in this review, in diverse contexts.
- MIND was initially tailored to mental health and thus does not encompass thorough disease-specific criteria for other conditions such as asthma, diabetes and sickle cell anaemia—though such questions may be easily integrated.
- Subjective questions, especially those around ease of use and visual appeal, are difficult to standardise but may be among the most important features driving user engagement with mental health apps.

select both an appropriate evaluation tool and to find an appropriate health app. This review provides a comparison of what different app evaluation frameworks are offering, where the field is converging and new priorities for improving clinical guidance.

INTRODUCTION

The past 5 years have seen a proliferation of both mobile health apps and proposed tools to rate such apps. While these digital health tools hold great potential, concerns around privacy, efficacy and credibility, coupled with a lack of strict oversight by governing bodies, have highlighted a need for frameworks that can help guide clinicians and consumers to make informed app choices. Although the US' Food and Drug Administration has recognised the issue and is piloting a precertification programme that would prioritise app safety at the developer level,¹ this model is still in pilot stages and there has yet to be an international consensus around standards for health apps, resulting in a profusion of proposed frameworks across governments, academic institutions and commercial interests.



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

Division of Digital Psychiatry, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts, USA

Correspondence to

Dr John Torous;
jtorous@bidmc.harvard.edu

In 2018, our team drew on existing evaluation frameworks to identify salient categories from existing rating schemes and create a new framework.² The American Psychiatric Association's (APA) App Evaluation Model was developed by harmonising questions from 45 evaluation frameworks and selecting 38 total questions that mapped to five categories: background information, privacy and security, clinical foundation, ease of use and interoperability. This APA model subsequently has been used by many diverse stakeholders given its flexibility in guiding informed decision-making.^{3–7} However, the flexibility of the model also created a demand for a more applied approach that offered users more concrete information instead of placing the onus entirely on a clinician or provider.

Thus, since the framework's development, the initial 38 questions have been operationalised into 105 new objective questions that invite a binary (yes/no) or numeric response by a rater.⁸ These questions align with the categories proposed by the APA model but are more extensive and objective, with, for example, 'app engagement' operationalised into 11 different engagement styles to select. These 105 questions are sorted into six categories (App Origin and Functionality, Inputs and Outputs, Privacy and Security, Clinical Foundation, Features and Engagement, Interoperability and Data Sharing) and are intended to be answerable for any trained rater—clinician, peer, end user—and inform the public-facing Mhealth Index and Navigation Database (MIND), where users can view app attributes and compare ratings (see figure 1 below). MIND, thus, constitutes a new framework

based on the APA model, with an accompanying public-facing database.

Recent systematic reviews have illustrated the growing number of evaluation tools for digital health devices, including mobile health apps.^{9–11} Given the rapidly evolving health app space and the need to understand what aspects are considered in evaluation frameworks, we have sought to survey the landscape of existing frameworks. Our goal was to compare the categories and questions composing other frameworks to (1) identify common elements between them, (2) identify if gaps in evaluation frameworks have improved since 2018 and (3) assess how reflective our team's MIND framework is in the current landscape. We, thus, aimed to map every question from the 2018 review, as well as questions from new app evaluation frameworks that have emerged since, using the questions of MIND as a reference. While informing our own efforts around MIND, the results of this review offer broad relevance across all of digital health, as understanding the current state of app evaluation helps inform how any new app may be assessed, categorised, judged and adopted.

METHODS

Patient and public involvement

Like the APA model, MIND shifts the app evaluation process away from finding one 'best' app, and instead guiding users towards an informed decision based on selecting and placing value on the clinically relevant criteria that account for the needs and preferences of

Figure 1 A screenshot of MIND highlighting several of the app evaluation questions (green boxes) and ability to access more. MIND, Mhealth Index and Navigation Database.

each patient and case. Questions were created with input of clinicians, patients, family members, researchers and policy-makers. The goal is not for a patient or clinician to consider all 105 questions but rather be able to access a subset of questions that appear most appropriate for the current use case at hand. Thus, thanks to its composition of discrete questions that aim to be objective and reproducible, MIND offers a useful tool to compare evaluation frameworks. It also offers an actionable resource for any user anywhere in the world to engage with app evaluation, providing tangible results in the often more theoretical world of app evaluation.

Design

We followed a three-step process in order to identify and compare frameworks to MIND. This process included (1) assembling all existing frameworks for mobile medical applications, (2) separating each framework into the discrete evaluation questions comprising it and (3) mapping all questions to the 105 MIND framework questions as a reference.

Search strategy and selection criteria

We started with the 45 frameworks identified in the 2018 review by Moshi *et al*⁹ and included 34 frameworks that have emerged since our initial analysis of the space that was conducted in 2018 and published in 2019.² To accomplish this, we conducted an adapted scoping review based on the Moshi criteria to identify recent frameworks. Although MIND focuses on mental health apps, its considerations and categories are transferable to health apps more broadly, and, thus, there was no mental health specification in the search terms.

References were identified through searches of PubMed, EMBASE and PsychINFO with the search terms ((mobile application) OR (smartphone app)) AND ((framework) OR (criteria) OR (rating)) and publication date between January 2018 and October 2020. We also identified records beyond the database search by seeking frameworks mentioned in subsequent and recent reviews^{5 12 13} and surveying the grey literature and government websites. Papers were selected for inclusion if they meet the predetermined eligibility criteria—presenting an evaluation framework for mobile health apps with patient, clinician or end user-facing questions. Two reviewers (SL and JT) screened the literature separately and applied the inclusion criteria. The data extracted from the papers included: author and dates of publication, source affiliation, country of origin, name of framework, study design, description of framework, intended audience/user and framework scoring system. Articles were screened if they describe the evaluation of a single app, did not present a new framework (instead conducting a review of the space or relying on a previous framework), the framework was focused on developer instead of clinicians or end users, was the implementation and not evaluation focused, was not a framework for health apps and was a satisfaction

survey instead of an evaluation framework. The data selection process is outlined in [figure 2](#).

The 34 frameworks identified in the search were combined with the 45 frameworks from the 2018 review for a total of 79 frameworks for consideration. To our knowledge, this list comprehensively reflects the state of the field at the time of assembly. However, we do not claim it to be exhaustive, as frameworks are constantly changing, emerging and sunsetting, with no central repository. The final list of frameworks assembled can be found in online supplemental appendix 1.

Mapping

Each resulting framework was reviewed and compiled into a complete list of its unique questions. The 79 frameworks yielded 1701 questions in total. Several of the original 45 frameworks focused exclusively on in-depth privacy considerations (evaluating privacy and security practices rather than the app itself),¹⁴ and after eliminating these checklists that did not facilitate app evaluation by a clinician or end user, 70 total frameworks were mapped in entirety to the MIND framework.

In mapping questions, discussion was sometimes necessary as not every question was an exact, word-for-word match. The authors, thus, used discretion when it came to matching questions to MIND and discussed each decision to confirm mapping placement. Two raters (SL, LS) agreed on mapping placement, and disputes were brought to a third reviewer (JT) for final consideration. ‘Is data portable and interoperable?’¹⁵ for example, would be mapped to the question ‘can you email or export your data?’ ‘Connectivity’¹⁶ was mapped to ‘Does the app work offline?’ and ‘Is the arrangement and size of buttons/content on the screen zoomable if needed’¹⁷ was mapped to ‘is there at least one accessibility feature?’ Questions about suitability for the ‘target audience’ were mapped to the ‘patient-facing’ question in MIND.

RESULTS

Framework type

The aim of this review was to identify and compare mobile health app rating frameworks, assessing overlap and exploring changes and gaps relative to both previous reviews and to the MIND framework. Of the 70 frameworks ultimately assessed and mapped, the majority 39 (55.7%) offered models for evaluating mobile health apps broadly. Seven (10%) considered mental health apps, while six (8.5%) focused on apps for diabetes management. Other evaluation focuses included apps for asthma, autism, concussions, COVID-19, dermatology, eating disorders, heart failure, HIV, pain management, infertility and sickle cell disease ([table 1](#)).

Mapping

We mapped questions from 70 app evaluation frameworks against the six categories and 105 questions of MIND (see online supplemental appendix 2). We examined

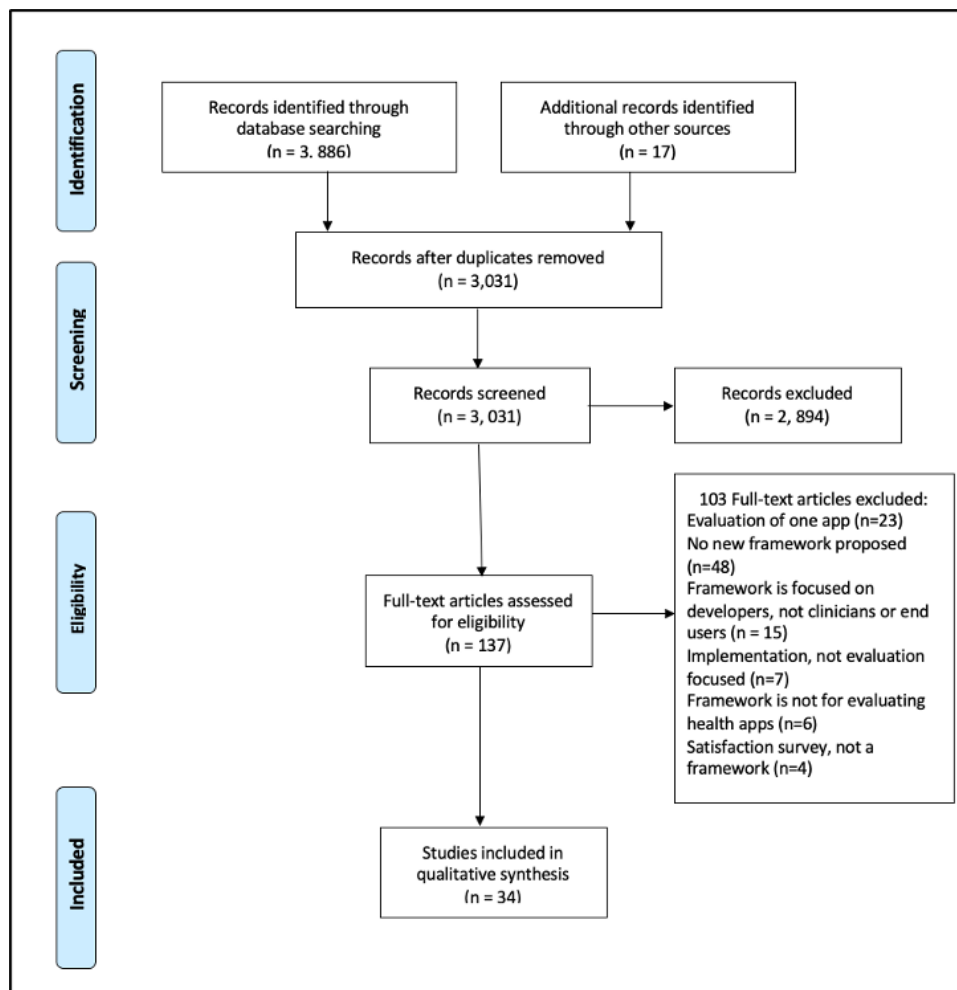


Figure 2 Framework identification through database searches (PubMed, EMBASE, PsychINFO) and other sources (reviews since 2018, grey literature, government websites).

the number of frameworks that addressed each specific MIND category and identified areas of evaluation that are not addressed by MIND. Through the mapping process, we were able to gauge the most common questions and categories across different app evaluation frameworks.

We sorted the questions into MIND's six different categories—App Origin & Functionality, Inputs & Outputs, Privacy & Security, Evidence & Clinical Foundation, Features & Engagement Style and Interoperability & Data Sharing—in order to assess the most common broad areas of consideration. Across frameworks, the most common considerations were around privacy/security and clinical foundation, with 43 frameworks posing at least one question around the app's privacy protections and 57 of the frameworks containing at least one question to evaluate evidence base or clinical foundation, as denoted in [table 2](#). Fifty-nine frameworks covered at least two of the MIND categories, with the majority of frameworks overlapping with at least four of MIND categories.

We then took a more granular look at the questions from each of the 70 frameworks, matching questions one-by-one to questions of the MIND framework when possible. On an individual question level, specific

questions about the presence of a privacy policy, security measures in place, supporting studies and patient-facing (or target population) tools were the most prevalent, with representation from 20, 25, 27 and 28 frameworks, respectively, for each question. Each of the 70 frameworks had at least one question that mapped to MIND. The most common questions, sorted into their respective categories, are depicted in [figure 3](#) and [table 3](#), while the full list of mapped questions can be found in online supplemental appendix 2.

Every question was examined but not every question in every framework could be matched to a corresponding question in MIND, and some questions fell outside one of the six categories. For example, 18 frameworks continue to present the subjective question of 'is the app easy to use' which will vary depending on the person and use case. MIND also does not offer questions related to other objective questions to which answers are not readily available such as 'How were target users involved in the initial design and usability evaluations of the app?'¹⁸ While questions such as this are of high importance, lack of easily accessible answers creates a dilemma in their present utility for app evaluation. Furthermore, some questions

Table 1 Number of disease-specific and general app evaluation frameworks, with general mobile health frameworks constituting more than half of identified frameworks

Focus of evaluation	Number of frameworks
Asthma	2 (2.86%)
Autism	1 (1.43%)
Concussion	1 (1.43%)
COVID-19	1 (1.43%)
Dermatology	1 (1.43%)
Diabetes	6 (8.57%)
Eating disorders	1 (1.43%)
General mobile health	39 (55.71%)
Heart failure/cardiac rehabilitation	3 (4.29%)
HIV	2 (2.86%)
Mental health	7 (10%)
Pain management	2 (2.86%)
Paediatric or adolescent health	2 (2.86%)
Reproductive endocrinology/infertility	1 (1.43%)
Sickle cell disease	1 (1.43%)

such as economic analysis were not covered by MIND but by other frameworks and represent a similar dilemma in that actual data to base evaluation on are often lacking. Aside from subjective questions, other pronounced absences MIND were questions about customisability (addressed by seven other frameworks) and advertising (nine frameworks). Although MIND does ask about customisability in part by encouraging raters to consider accessibility features (and some frameworks ask about the ability to customise in conjunction with accessibility features,¹⁹ MIND neither pose a question around the user's ability to tailor or customise app content nor does it ask questions about the presence of advertisements on an app. Other questions unaddressed by MIND were about

Table 2 The questions from all frameworks were mapped to the reference framework (MIND) sorted into its six categories, with this table denoting how many frameworks had questions that could be sorted into each of the categories

Category of evaluation	Number of frameworks addressing category
App origin/app functionality	42
Inputs and outputs	17
Privacy/security	43
Evidence/clinical foundation	57
Features/engagement style	29
Interoperability and data sharing	23

MIND, Mhealth Index and Navigation Database.

the user's ability to contact the producer or developer to seek guidance about app use. Variations of this question include 'is there a way to feedback user comments to the app developer?' MIND also does not pose any questions regarding instructions in the app or the existence of a user guide.²⁰ Finally, it does ask about speed of app functionality. This variant of question asks, 'is the app fast and easy to use in clinical settings?'¹⁵ figure 3 above, and table 3 below presents additional details on categories and questions both inside and outside the MIND reference framework.

DISCUSSION

As mobile health apps have proliferated, choosing the right one has become increasingly challenging for patient and clinician alike. While app evaluation frameworks can help sort through the myriad of mobile health apps, the growing number of frameworks further complicates the process of evaluation. Our review examined the largest number of evaluation frameworks to date with the goal of assessing their unique characteristics, gaps as well as overlap with the 105 questions in MIND. We identified frameworks for evaluating a wide range of mobile health apps—some focused on general mobile health, some specific and addressing specific disease domains like asthma, heart failure, mental health or pain management.

Despite the different disease conditions they addressed, there was substantial overlap among the frameworks, especially around clinical foundation and privacy and security. The most common category addressed was clinical foundation, with 57 of the evaluation frameworks posing at least one question regarding evidence base. More than half of the frameworks also addressed privacy and/or security and app functionality or origin.

The widespread focus on clinical foundation and privacy represents a major change in the space since 2018, when our team analysed an initial review of 45 health app evaluation frameworks and found that the most common category of consideration among the different frameworks was usability, with short-term usability highly over-represented compared with privacy and evidence with base. In this 2018 review, there were 93 unique questions corresponding to short-term usability but only 10 to the presence of a privacy policy. Although many frameworks continue to consider usability, our current review suggests the most common questions across frameworks now concern evidence, clinical foundation and privacy. This shift may reflect an increased recognition of the privacy dangers some apps may pose.

This review illustrates the challenges in conceiving a comprehensive evaluation model. A continued concern in mobile health apps is engagement,⁶ and it is unclear whether any framework adequately predicts engagement. Another persistent challenge is striking a balance between transparency/objectivity and subjectivity. Questions that prompt consideration of subjective user experiences may limit the generalisability and standardisation

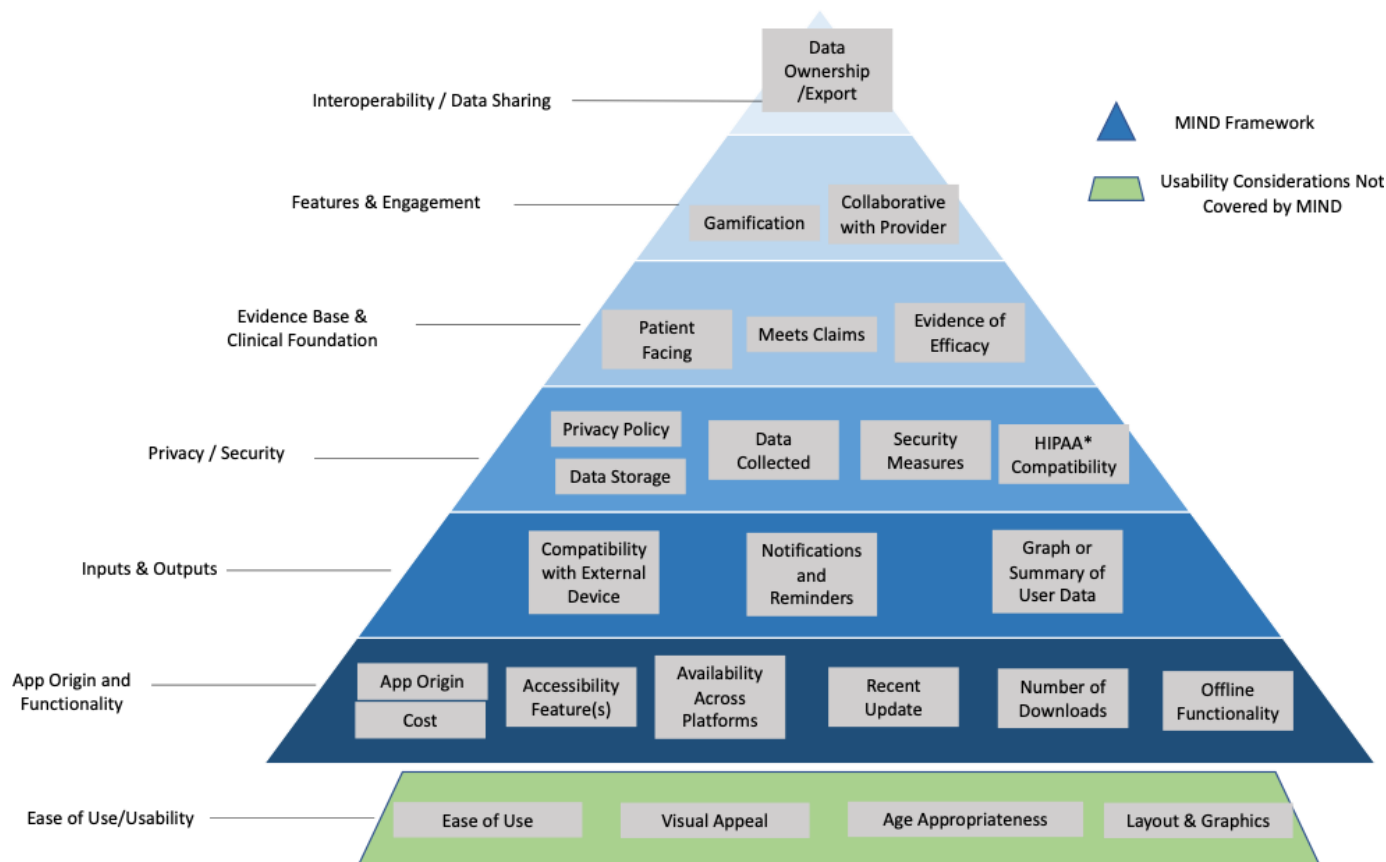


Figure 3 The most commonly addressed questions, grouped within the categories of MIND. The blue triangle constitutes MIND and its six main categories, while the green trapezoid represents questions pertaining to usability or ease of use, which are not covered by MIND. MIND, Mhealth Index and Navigation Database.

of a framework, as the questions inherently reflect the experience of the rater. An app's ease of use, for example, will differ significantly depending on an individual's level of comfort and experience with technology. However, subjective questions around user friendliness, visual appeal and interface design may be of greatest concern to an app user, and most predictive of engagement with an app.²¹ Finally, a thorough assessment of an app is only feasible if information about the app is available. For example, some questions with clinical significance, such as the consideration of how peers or target users may be involved in app development, are not easily answerable by a health app consumer. Overall, there is a need for more data and transparency when it comes to health apps. App evaluation frameworks, while thorough, rigorous and tailored to clinical app use, can only go so far without transparency on the part of app developers.²²

The analysis additionally highlighted the flexibility and comprehensiveness of the MIND framework, which was used as a reference framework in this review, in diverse contexts. The MIND categories are inclusive of a wide range of frameworks and questions. Even without including any subjective questions in the mapping process, each of the 70 frameworks that were ultimately mapped had some overlap with MIND, and many of the 1700 questions ultimately included were mapped exactly with a MIND question. Although MIND was initially

conceptualised as an evaluation tool specifically for mental health apps, the coherence between MIND and diverse types of app evaluation frameworks, such as those for concussion,²³ heart disease²⁴ and sickle cell anaemia,²⁵ demonstrates how the MIND categories can encompass many health domains. Condition-specific questions, for example, are a good fit for the 'Features & Engagement' category of MIND.

The results of our analysis suggest while numerous new app evaluation frameworks continue to emerge, there is a naturally appearing standard of common questions asked across all. While different use cases and medical subspecialties will require unique questions to evaluate apps, there are a set of common questions around aspects like privacy and level of evidence that are more universal. MIND appears to cover a large subset of these questions and, thus, may offer a useful starting point for new efforts as well as means to consolidate existing efforts. Advantage of the more objective approach offered by MIND is that it can be represented as a research database to facilitate discovery of apps while not conflicting with local needs, personal preferences or cultural priorities.²⁶

Limitations

Our work is not the first to compare app evaluation frameworks. Recently, several reviews have compared how different mobile health app evaluation models address

Table 3 Commonly addressed questions among those that could be mapped to the MIND reference framework (blue), and those that could not (green)

Question	Number of frameworks addressing question
App origin/app functionality: Where does the app come from/who is the app developer?	26
App origin/app functionality: How much does the app cost?	16
App origin/app functionality: Does the app have at least one accessibility feature?	11
App origin/app functionality: Is the app available on multiple platforms?	9
App origin/app functionality: Has the app been updated recently/when was the last update?	9
App origin/app functionality: How many downloads does the app have/how many reviews?	9
App origin/app functionality: Does the app work offline?	7
Privacy/security: Is there a privacy policy?	20
Privacy/security: Does the app declare data use and purpose?	18
Privacy/security: Does the app report security measures in place?	25
Privacy/security: Does the app claim to meet HIPAA (or an analogous health data protection regulation)?	10
Evidence/clinical foundation: Is the app content well-written, correct and relevant?	25
Evidence/clinical foundation: Does the app do what it claims?	18
Evidence/clinical foundation: Is the app patient facing?	28
Evidence/clinical foundation: Is there evidence of the app's efficacy?	27
Features/engagement style: Is the app collaborative with a provider?	8
Features/engagement style: Does the app offer gamification?	4
Interoperability/data sharing: Can you email or export your data?	10
Is the app easy to use?	18
Is the app visually appealing?	10
Is the app age appropriate?	4
Does the app allow for customisation and/or personalisation?	7
Was the app developed in collaborated with target users?	4
Is there a manual or set of instructions for the app?	8
Is there any advertising within the app?	9
Risk/benefit analysis	2
Speed of functionality	5
Is there a way to contact the developer with questions or concerns?	6

HIPAA, Health Insurance Portability and Accountability Act; MIND, Mhealth Index and Navigation Database.

privacy,^{11 12 14} and another database (<https://search.appcensus.io/>) focuses exclusively on compiling privacy assessments of Android apps. We chose to exclude app evaluation frameworks that focused exclusively on in-depth privacy considerations and were unusable by a clinician or layperson as our goal was more comprehensive app evaluation. This decision is not to reject considerations of privacy and security that are of critical importance, but rather to narrow the focus to frameworks that are usable in the hands of the public today and can be used to inform clinical decisions. In addition, MIND was initially tailored to mental health, and thus does not encompass thorough disease-specific criteria for other conditions such as asthma, diabetes and sickle cell anaemia—though such questions may be easily integrated. Finally, subjective

questions, especially those around ease of use and visual appeal, are difficult to standardise but may be among the most important features driving user engagement with mental health apps.²¹

CONCLUSION

Our work demonstrates the expansion of app evaluation frameworks. By illustrating how the MIND overlaps with many of these existing and emerging frameworks—we suggest the practical need for consolidation. Although specific disease tailored mobile health apps require specialised app evaluation *questions*, concerns around accessibility, privacy, clinical foundation and interoperability are nonspecific. If the full potential of digital health

can be realised, there is a need for increased collaboration among industry, government and academia in order to ensure that the highest quality digital health tools reach the public. We emphasise that this effort is just a first step and highlight the need for interdisciplinary continued communication among diverse digital health stakeholders in order to best serve the public.

Contributors SL and JT designed the procedure. SL and JT screened articles for eligibility. SL and LS compiled and mapped questions from frameworks. SL and JT composed manuscript.

Funding This work was supported by a gift from the Argosy Foundation.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Additional data are presented in Appendix A and B.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

John Torous <http://orcid.org/0000-0002-5362-7937>

REFERENCES

- 1 FDA. Health C for D and R. digital health policies and public health solutions for COVID-19, 2020. Available: <https://www.fda.gov/medical-devices/coronavirus-covid-19-and-medical-devices/digital-health-policies-and-public-health-solutions-covid-19>
- 2 Henson P, David G, Albright K, *et al*. Deriving a practical framework for the evaluation of health apps. *Lancet Digit Health* 2019;1:e52–4.
- 3 Ondersma SJ, Walters ST. Clinician's Guide to Evaluating and Developing eHealth Interventions for Mental Health. *Psychiatr Res Clin Pract* 2020;2:26–33.
- 4 Bergin A, Davies EB. Technology Matters: Mental health apps - separating the wheat from the chaff. *Child Adolesc Ment Health* 2020;25:51–3.
- 5 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;17:231.
- 6 Carlo AD, Hosseini Ghomi R, Renn BN, *et al*. By the numbers: ratings and utilization of behavioral health mobile applications. *NPJ Digit Med* 2019;2:54.
- 7 Magee JC, Adut S, Brazill K, *et al*. Mobile APP tools for identifying and managing mental health disorders in primary care. *Curr Treat Options Psychiatry* 2018;5:345–62.
- 8 Lagan S, Aquino P, Emerson MR, *et al*. Actionable health APP evaluation: translating expert frameworks into objective metrics. *npj Digital Medicine* 2020;3:1–8.
- 9 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;34:464–75.
- 10 Coravos A, Doerr M, Goldsack J. Modernizing and designing evaluation frameworks for connected sensor technologies in medicine. *npj Digital Medicine* 2020;3:1–10.
- 11 Nurgalieva L, O'Callaghan D, Doherty G. Security and privacy of mHealth applications: a scoping review. *IEEE Access* 2020;8:104247–68.
- 12 Benjumea J, Roper J, Rivera-Romero O, *et al*. Privacy assessment in mobile health Apps: Scoping review. *JMIR mHealth and uHealth* 2020;8:e18868.
- 13 Levine DM, Co Z, Newmark LP, *et al*. Design and testing of a mobile health application rating tool. *npj Digital Medicine* 2020;3:1–7.
- 14 Hussain M, Zaidan AA, Zidan BB, *et al*. Conceptual framework for the security of mobile health applications on android platform. *Telematics and Informatics* 2018;35:1335–54.
- 15 McMillan B, Hickey E, Patel MG, *et al*. Quality assessment of a sample of mobile app-based health behavior change interventions using a tool based on the National Institute of health and care excellence behavior change guidance. *Patient Educ Couns* 2016;99:429–35.
- 16 Shaia KL, Farag S, Chyjek K, *et al*. An evaluation of mobile applications for reproductive endocrinology and infertility providers. *Telemed J E Health* 2017;23:254–8.
- 17 Stoyanov SR, Hides L, Kavanagh DJ, *et al*. Mobile APP rating scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015;3:e27.
- 18 Wykes T, Schueller S. Why reviewing Apps is not enough: transparency for trust (T4T) principles of responsible health APP marketplaces. *J Med Internet Res* 2019;21:e12390.
- 19 Nebeker C, Bartlett Ellis RJ, Torous J. Development of a decision-making checklist tool to support technology selection in digital health research. *Transl Behav Med* 2020;10:1004–15.
- 20 Noee M, Akbari Sari A, Olyaeemanesh A, *et al*. Prioritizing the potential applications of Mobile-Health in the Iranian health system. *J Res Health Sci* 2020;20:e00473.
- 21 Alqahtani F, Orji R. Insights from user reviews to improve mental health apps. *Health Informatics J* 2020;26:2042–66.
- 22 Rodriguez-Villa E, Torous J. Regulating digital health technologies with transparency: the case for dynamic and multi-stakeholder evaluation. *BMC Med* 2019;17:226.
- 23 Lee H, Sullivan SJ, Schneiders AG, *et al*. Smartphone and tablet apps for concussion road warriors (team clinicians): a systematic review for practical users. *Br J Sports Med* 2015;49:499–505.
- 24 Farnia T, Jalent M-C, Steichen O. Evaluation criteria of noninvasive Telemonitoring for patients with heart failure: systematic review. *J Med Internet Res* 2018;20:e16.
- 25 Shah N, Jonassaint J, De Castro L. Patients welcome the sickle cell disease mobile application to record symptoms via technology (smart). *Hemoglobin* 2014;38:99–103.
- 26 Hoffman L, Wisniewski H, Hays R, *et al*. Digital opportunities for outcomes in recovery services (doors): a pragmatic Hands-On group approach toward increasing digital health and smartphone competencies, autonomy, relatedness, and alliance for those with serious mental illness. *J Psychiatr Pract* 2020;26:80–8.