



Digital Health and Engagement—Looking Behind the Measures and Methods

John Torous, MD; Erin E. Michalak, PhD; Heather L. O'Brien, PhD

Low real-world engagement with digital mental health therapy programs and apps is a perennial concern that is only now receiving due research attention. The study by Chien et al¹ is a welcome addition to the literature given its focus on understanding the association between patient engagement, an internet-based cognitive behavioral therapy intervention for depression and anxiety, and clinical outcomes. Using a probabilistic latent variable model to infer distinct subtypes of patients based on their interaction with the internet-based cognitive behavioral therapy program, the authors identify 5 clusters of patients: low engagers, late engagers, high engagers with rapid disengagement, high engagers with moderate decrease, and highest engagers. Overall, the results are consistent with the literature on digital health engagement: highest engagers represented just 10.6% of the sample.

But low engagement measurements with digital health tools do not tell the full story. The study is obfuscated not by the technology (most digital therapy apps have similar engagement patterns as reported in this article) but by the measurement of engagement itself. Defining meaningful engagement, or even feasibility, for digital mental health tools remains a challenge for the field and comparing engagement across studies or products remains nearly impossible.² In this study, Chien et al¹ assessed 2 broad types of engagement, specifically whether, in a given week, the patient used the platform or particular sections of the platform. These user engagement measurements were typical of this nascent field (ie, behavioral) and therefore relatively blunt. This potential limitation is noted by the authors, who reflect that raw counts of use cannot account for the complexity or diversity in the type of content that is embedded in them.

Behavioral measures tend to capture product rather than process. A more-nuanced understanding of user engagement recognizes that engagement is a process and a product of people's interactions with digital tools. Degree of engagement is influenced by the depth of the patient's investment in the interaction with the digital tool; this investment may be defined temporally, affectively, and/or cognitively.³ Behavioral measures represent one important measure, but do not effectively capture affective and cognitive investment. To understand where people are starting from emotionally and cognitively is particularly important in mental health. In a patient experiencing more severe symptoms, it can be a major accomplishment to simply login into an app and connect with a support person, never mind explore interactive features, such as quizzes, read content, and journal. For some people, this single action itself might be highly influential, especially if the program or app is recovery focused. Other users may benefit from frequent, sustained interactions with the app that allow them to better understand and monitor their wellness needs and triggers. In both of these instances, each user "gets what they need" from the technology, but their patterns of engagement look very different. Although it is essential to take a longitudinal perspective with engagement and examine use over time, duration of use itself is not a reliable indicator of engagement. Studies in human-computer interaction have shown that it is difficult to disambiguate negative, frustrating experiences with technology from positive, absorbing ones based on this measure, and that a person's willingness to return is more telling of engagement.⁴ Thus, we might ask whether the user is continuing to come back to the app rather than focusing on their session length and degree of active engagement with different tools. Categorizing users can be helpful for understanding that peoples' engagement trajectories with applications vary based on their individual motivations, goals, and needs, but classifying users can also paint them as quite static rather than responding to dynamic shifts in managing their mental health.

+ Related article

Author affiliations and article information are listed at the end of this article.

Open Access. This is an open access article distributed under the terms of the CC-BY License.

Chien et al¹ recognize this limitation in their observation that time in the platform was not necessarily indicative of clinical outcomes. User engagement likely operates along a continuum from shallow to deep, and this continuum will be influenced not only by the nature of the interaction but also contextual constraints and users' mental health goals and needs. Behavioral measures do not necessarily allow us to fully appreciate this continuum or distinguish shallow and game-like from deep and meaningful engagement.

Clinically, interpreting these seemingly simple results now becomes more complex. One notable finding from the study is both floor and dose effects for program users. Regardless of engagement cluster, all patients showed some clinical improvements. Whether a digital placebo effect, regression to the mean, or other factors,⁵ these results remind us that a control group is vital to fully understand the outcome associated with any intervention. The observed dose effect (with high engagers experiencing more clinical improvement compared with lower engagement users) is suggestive that actual therapeutic mechanisms are offered by the program. Teasing apart this therapeutic association from confounding variables (eg, the role of the human supporter) remains a broad question as the field looks to increase scalability of software. Of course, understanding what an effective dose of an app is and how it varies by each person remains unknown. In the future, using static benchmarks to evaluate ongoing engagement with the intervention,⁶ as well as offering adaptive content, may offer a path to more effective and engaging digital mental health tools.

Artificial intelligence and machine learning are often proposed methods to advance this next generation of digital mental health tools. Chien et al's¹ use of machine learning focuses on different tasks but still serves as a useful example of advantages and challenges. Any machine learning approach in health care is always of interest, but also deserves careful interrogation around assumptions and replications. Digital health platforms are not static and continually evolve, and the pooled analysis of data from 2015-2019 offers advantages in terms of sample size but not the outcomes of different versions of the software. Offering analysis of 54 604 patients using one module of the system is impressive but also begs the question of whether replication efforts could immediately be undertaken on other therapy in the software modules. The prospective predictive power of these classifications to inform care remains the untested potential; a recent review on machine learning for suicide prediction offered the sobering statistic that even high classification accuracy still only resulted in a positive predictive value of under 1%.⁷

User engagement with digital mental health is the new primary challenge in realizing the full potential of scalable and accessible care. Thus, it is encouraging to see companies studying their own products and sharing data to advance the field, regardless of whether this process affects their profit margins. The next frontiers in this work must, however, look inward and better define and measure engagement. Only then will new tools, such as machine learning, advance the field outward and truly promote engaging and effective care.

ARTICLE INFORMATION

Published: July 17, 2020. doi:[10.1001/jamanetworkopen.2020.10918](https://doi.org/10.1001/jamanetworkopen.2020.10918)

Open Access: This is an open access article distributed under the terms of the [CC-BY License](https://creativecommons.org/licenses/by/4.0/). © 2020 Torous J et al. *JAMA Network Open*.

Corresponding Author: John Torous, MD, Department of Psychiatry, Beth Israel Deaconess Medical Center, Harvard Medical School, 330 Brookline Ave, Boston, MA 02215 (jtorous@bidmc.harvard.edu).

Author Affiliations: Department of Psychiatry, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts (Torous); Department of Psychiatry, University of British Columbia, Vancouver, British Columbia, Canada (Michalak); School of Information, University of British Columbia, Vancouver, British Columbia, Canada (O'Brien).

Conflict of Interest Disclosures: Dr Torous reported receiving research support from Otsuka Pharmaceuticals outside the submitted work, a research grant from the National Institutes of Mental Health (1K23MH116130-01), a philanthropic grant from Jeremy Wertheimer. No other disclosures were reported.

REFERENCES

1. Chien I, Enrique A, Palacios J, et al. A machine learning approach to understanding patterns of engagement with internet-delivered mental health interventions. *JAMA Netw Open*. 2020;3(7):e2010791. doi:10.1001/jamanetworkopen.2020.10791
2. Ng MM, Firth J, Minen M, Torous J. User engagement in mental health apps: a review of measurement, reporting, and validity. *Psychiatr Serv*. 2019;70(7):538-544. doi:10.1176/appi.ps.201800519
3. O'Brien H. Theoretical perspectives on user engagement. In: *Why Engagement Matters*. Springer International Publishing; 2016:1-26.
4. O'Brien HL, Cairns P, Hall M. A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. *Int J Hum Comput Stud*. 2018;112:28-39. doi:10.1016/j.ijhcs.2018.01.004
5. Cohen J, Torous J. The potential of object-relations theory for improving engagement with health apps. *JAMA*. 2019;322(22):2169-2170. doi:10.1001/jama.2019.17141
6. Fletcher K, Foley F, Murray G. Web-based self-management programs for bipolar disorder: insights from the online, recovery-oriented bipolar individualised tool project. *J Med internet Res*. 2018;20(10):e11160. doi:10.2196/11160
7. Belsher BE, Smolenski DJ, Pruitt LD. Positive predictive values and potential success of suicide prediction models—reply. *JAMA Psychiatry*. 2019;76(8):870-871. doi:10.1001/jamapsychiatry.2019.1510