

Feature Article

# A Decade of Ubiquitous Computing Research in Mental Health

**Jakob E. Bardram**

Department of Health Technology, Technical University of Denmark

**Aleksandar Matic**

Telefonica Innovacion Alpha

**Abstract—Mental health represents a huge disease and societal burden and a significant body of research in ubiquitous computing has been devoted to the design of technologies for continuous monitoring, diagnosis, and care of mental health conditions. This paper reviews a decade of research into technologies for mental health, focusing on the use of mobile and wearable technology. The review found 46 systems that are analyzed in a historical context and discussed according to which mental disorder they target, the type of technology, and the type and size of the clinical studies they have been used in. Finally, the paper presents inputs from nine leading researchers in the domain and discuss important technical and clinical challenges in the design of ubiquitous computing technology for the next decade.**

■ **MENTAL HEALTH REPRESENTS** a huge disease and societal burden.<sup>1</sup> Due to its episodic nature, the traditional healthcare model is considered to be suboptimal to address this burden and to improve chronic mental conditions.<sup>2</sup> “Anytime and everywhere” ubiquitous technology was seen early on as an opportunity to address continuous monitoring, diagnosis, and care of mental health conditions, thereby enabling an extension of care

delivery beyond the reach of traditional health-care. In particular, mobile and wearable technologies—with their ability to track behavioral, physiological, and contextual signals—were seen as a potential enabler of a continuous symptom monitoring and personalized intervention.<sup>3,4</sup>

Last year marked the ten year anniversary since smartphones (the first iPhone appeared in 2007 and the first stable Android phone in 2009) have become widely available as an open platform and have since then been used for creating novel personalized health applications. The early

*Digital Object Identifier 10.1109/MPRV.2019.2925338*

mobile phones with sensing capabilities had already sparked inspiration of pioneering UbiComp researchers<sup>5,6</sup> who recognized the opportunity for improving mental healthcare. This was followed up by a growing interest in exploiting advantages of mobile and wearable technologies to unobtrusively sense and analyze human behavior, assess, and predict mental health status, and to deliver timely feedback and intervention.<sup>3</sup>

In this paper, we look back at the last decade of Ubicomp research in mental health. We focus on mobile and wearable technologies and systems, which have been built and evaluated in research studies. Initially, we present a review of 46 systems developed and published over the last ten years. We report which mental disorders they were designed for, as well as which technical features they included—namely sensing, prediction, intervention, and clinical assessment. Furthermore, we present the opinion of nine leading researchers in this field who we interviewed to gather their retrospective and prospective view on the status of UbiComp research for mental health. By combining their input with the conducted technological review, we discuss study reproducibility, technical issues, clinical evidence, and clinical adoption that emerged as the key challenges in this field. Finally, we provide an outlook for future research for the next decade.

## METHODS

In contrast to a traditional systematic literature review of published papers—which is a standard approach in medical sciences—this review focuses on research-based technologies and systems for mental health. Hence, the “unit” of the review is not a study but a technology or a system, which has been published in one or more papers.

We used a snowballing approach to identify the relevant work. Snowballing refers to using the reference list and citations of one selected paper to find related papers.<sup>7</sup> Using references and citations is referred to backward and forward snowballing, respectively. In this review, we identified an initial “seed” set of nine technologies to be the starting point (marked in bold in Table 1). These seed papers were carefully selected to represent different types of technologies, from different geographical regions, published in different venues,

and from different years. We also thoroughly reviewed the existing systematic review papers to identify relevant systems.<sup>8,9,10,11,12,13,14</sup> Once we compiled the full list of technologies, we contacted a set of leading researchers in the field (see Acknowledgements) and asked for verification of the review list and annotations.

The review was conducted by applying the following inclusion criteria: first, research-based technologies published in scientific peer-reviewed papers; second, mobile and wearable technology; third, technologies tested for severe mental illness (SMI) as defined by ICD-10 on “Mental and behavioral disorders”<sup>15</sup> including schizophrenia, affective/mood disorders (including depression and bipolar disorder), neurotic and stress-related disorders (including stress, Post-Trauma Stress Disorder (PTSD), phobia, and anxiety), disorders of psychological development (including autism and attention deficit hyperactivity disorder (ADHD)), aging-related mental disorders (Alzheimer’s, dementia), and substance abuse. We also reviewed technologies that focused on mental well-being, but that were evaluated with healthy subjects, and we report these systems separately. Each system was labeled with respect to:

- (1) the year when the system was first published,
- (2) disease classification (according to ICD-10) and specific disorder(s),
- (3) geographical region (US, EU, Asia, Australia),
- (4) technology topic and type of technology (mobile, wearable), and
- (5) the size of a clinical study in terms of number of participants ( $N$ ) and duration ( $T$ ) in days.

The review was done in late 2018 and early 2019.

## RESULTS

Table 1 shows the list of identified technologies represented as systems with a clear SMI focus (upper section of Table 1) and systems focusing on general mental well-being and/or evaluated only with healthy subjects (bottom section of Table 1). Within each category, systems are listed in chronological order based on when the first scientific paper describing the system was published, as indicated in the last column. A copy of this table with a list of all 50 scientific papers are

**Table 1. Ubiquitous computing systems and technologies in mental health. Systems used as seed are marked in bold. ICD-10 categorization: MOOD—Mood (affective) disorders; SCHIZ—Schizophrenia, schizotypal and delusional disorders; DEV—Disorders of psychological development; NEU – Neurotic, stress-related and somatoform disorders; SUB—Mental and behavioral disorders due to psychoactive substance use; PDIS—Disorders of adult personality and behavior.**

System (N=46)	Year	ICD10 SMI	Disorder(s)	Region	Topics	Technology	Study (N/T)
<i>Clinical focus ( N=33 )</i>							
PsychLog	2010	NEU	Stress, PTSD	EU	SEN;INT	WEAR;MOB	100 / 270
LifeShirt	2010	SCHIZ;MOOD	Schizophrenia, Bipolar	US	SEN	WEAR	28 / 1
Empath	2011	MOOD	Depression	US	SEN	MOB	1 / 14
<b>Mobilyze!</b>	2011	MOOD	Depression	US	SEN;PRE;CAS	MOB	7 / 56
A-CHESS	2011	SUB	Alcohol abuse	US	SEN;INT	MOB	280 / 365
<b>MONARCA</b>	2012	MOOD	Bipolar	EU	UI;SEN;PRE;ASS	MOB	12 / 98
Moodbuster	2012	MOOD	Depression	EU	UI;INT;CAS	WEAR	52 / 30
MOSOCO	2012	DEV	Autism	US	INT;UI	MOB	12 / 49
AGATE	2012	SUB	Alcohol abuse	US	UI;INT	MOB	105 / 56
<b>StudentLife</b>	2013	MOOD	Depression	US	SEN;CAS	MOB	48 / 70
<b>MoodRhythm</b>	2013	MOOD	Bipolar	US	SEN;INT;UI	MOB	7 / 28
MONARCA <sup>2</sup>	2014	MOOD	Bipolar	EU	SEN;PRE	MOB	12 / 48
BigBlackDog	2014	MOOD	Depression	US	SEN	MOB	3 / 120
ParentGuardian	2014	DEV	ADHD	US	SEN;UI	MOB	10 / 14
PSYCHE	2015	MOOD	Bipolar	EU	SEN	WEAR;MOB	26 / 1
<b>PurpleRobot</b>	2015	MOOD	Depression	US	SEN;UI;CAS	MOB	18 / 14
MoodTraces	2015	MOOD	Depression	EU	SEN;PRE;CAS	MOB	28 / 14
Dem@Care	2015	ORG	Dementia	EU	SEN;UI	WEA	n/a / n/a
Mindful Moods	2015	MOOD	Depression	US	CAS;SEN	MOB	13 / 30
LifeRhythm	2016	MOOD	Depression	US	SEN;PRE;CAS	MOB	79 / n/a
SIMBA	2016	MOOD	Bipolar	US	SEN;PRE;CAS	MOB	13 / 360
AMoSS	2016	MOOD	Bipolar	EU	SEN;CAS	MOB;WEAR	50 / n/a
NEVERMIND	2016	MOOD	Depression	EU	SEN;UI	WEAR;MOB	15 / n/a
MOOS	2016	MOOD	Depression	EU	SEN;INT;CAS	MOB	126 / 14
CrossCheck	2016	SCHIZ	Schizophrenia	US	SEN;PRE	MOB	21 / 180
SleepSight	2016	SCHIZ	Schizophrenia	EU	SEN	MOB;WEAR	16 / 56
<b>MOBERO</b>	2016	DEV	ADHD	EU	UI;CAS	MOB	13 / 14
DEMOS	2016	MOOD;NEU	Depression, Anxiety	US	SEN;CAS	MOB;WEAR	72 / n/a
MedLink	2016	MOOD	Depression	US	UI;CAS;INT	MOB	11 / 56
<b>PocketSkill</b>	2018	DEV;MOOD; NEU;PDIS	Depression, Anxiety, Borderline, PTSD, Bipolar	US	UI;INT;CAS	MOB	73 / 28
BEIWE	2018	SCHIZ	Schizophrenia	US	SEN;PRE	MOB	17 / 90
BiAffect	2018	MOOD	Bipolar	US	SEN;PRE	MOB	8 / 19
RADAR	2018	MOOD	Depression	EU	UI;SEN	MOB;WEAR	66 / n/a
<i>Non-clinical focus ( N=13 )</i>							
<b>EmotionSense</b>	2010	N/A	Emotional state	EU	SEN;PRE	MOB	18 / 10
MoodSense	2011	N/A	'Daily mood'	ASIA	SEN;PRE	MOB	25 / 30
Mobile Sensing Platform	2011	MOOD	Depression	US	SEN	WEAR	8 / 10
StressSense	2012	NEU	Stress	US	SEN;PRE	MOB	14 / 3
MoodMiner	2012	N/A	'Daily mood'	ASIA	SEN	MOB	15 / 30
AMON	2012	N/A	Emotion, Mood, Stress	US	SEN	MOB	n/a / n/a
Funf, Affectiva, MotionLogger	2015	N/A	Sleep, Stress, Mental Health	US	SEN	MOB;WEAR	66 / 30
MoodPrism	2016	N/A	'Emotional well-being'	AUS	SEN;UI	MOB	11 / 30
iHOPE	2016	MOOD;NEU	Depres., Anxiety, Stress	ASIA	SEN;PRE	MOB	28 / 14
Fine	2016	MOOD	Depression	EU	SEN	MOB	4 / 7
<b>EmotiCal</b>	2017	N/A	'Daily mood'	US	SEN;UI;PRE;INT	MOB	60 / 21
SHIM	2017	N/A	'Mental well-being'	EU	INT	MOB	28 / 14
AWEAR	2018	DEV	Dementia	EU	SEN	MOB;WEAR	5 / 7

available as supplementary material available at the IEEE PERSVASIVE COMPUTING MAGAZINE home page.

This review yielded a set of findings and observations that characterize trends in Ubi-Comp research in mental health over the last decade. First, in total 46 systems were identified of which 33 had a clear SMI focus and were subject to clinical evaluation, whereas 13 systems

had a non-SMI focus (labeled as “N/A” in Table 1) or were not subject to clinical evaluation. Researchers in this domain have often argued about the difficulty to conduct a technology research while also ensuring an appropriate clinical evaluation,<sup>16,3</sup>—therefore, it was encouraging to witness that the majority of the systems (72%) were clinically evaluated.

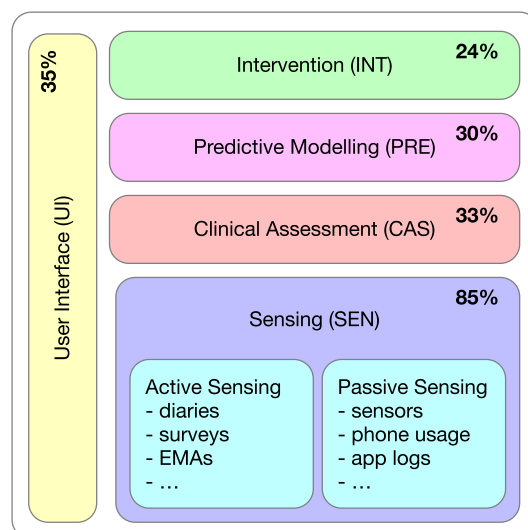
**Table 2. SMI focus of systems according to ICD-10. Since some systems supports more than one ICD-10 code, the sum is greater than 100%.**

Label	ICD-10		N	%
ORG	F00-F09	Organic, including symptomatic, mental disorders	1	2%
SUB	F10-F19	Mental and behavioural disorders due to psychoactive substance use	2	4%
SCHIZ	F20-F29	Schizophrenia, schizotypal and delusional disorders	4	9%
MOOD	F30-F39	Mood (affective) disorders	26	57%
NEU	F40-F48	Neurotic, stress-related and somatoform disorders	5	11%
BEH	F50-F59	Behavioural syndromes associated with physiological disturbances and physical factors	0	0%
PDIS	F60-F69	Disorders of adult personality and behaviour	1	2%
RETA	F70-F79	Mental retardation	0	0%
DEV	F80-F89	Disorders of psychological development	5	11%
CHLD	F90-F98	Behavioural and emotional disorders with onset in childhood and adolescence	0	0%
N/A		Unspecified mental disorder or healthy subjects	8	17%

Second, we analyzed which SMI disorders the systems focuses on—Table 2 shows the distribution according to the ICD-10 classification. We observed that the majority of systems were designed for mood (affective) disorders including depression and bipolar disorder (57%). Systems in this group include Moodbooster and Mobilize! (that focused on depression), and MONARCA, MoodRhythm, and BiAffect (focused on bipolar disorder). The three disorders classifications, namely first, schizophrenia, second, neurotic and stress-related disorders, and third, disorders of psychological development (e.g., Autism and ADHD), were each covered with a similar number of systems (9%–11%). A very few systems addressed substance abuse, organic disorders (e.g., Alzheimer’s, dementia), and disorders of adult personality (e.g., borderline). No systems addressed behavioral syndromes (e.g., eating or sleeping disorders), mental retardation, or childhood and adolescence disorders.

Third, by using a method of thematic labeling, the review revealed five types of technological “topics” addressed by different systems.

- *Sensing (SEN)*: technologies focusing on collection of data either from mobile phone or from wearable sensors, feature extraction from such data, as well as actively collected

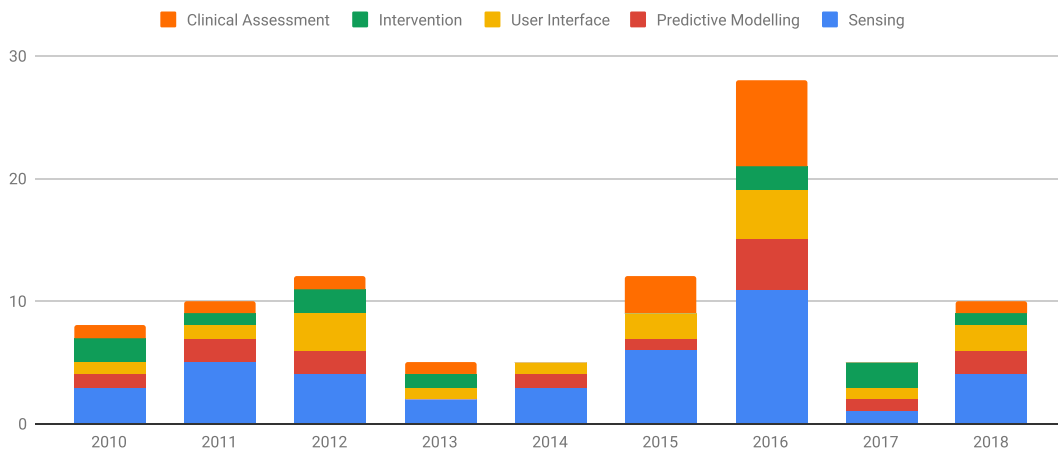


**Figure 1.** Technology topics and the number of systems focusing on each technology in percentage. Since most systems focus on more than one technology, the sum is greater than 100%.

data from users in form of surveys, questionnaires, and ecological momentary assessments (EMAs).

- *Clinical assessment (CAS)*: technologies for clinical assessment of disorder stages and symptoms, including questionnaires (e.g., PHQ9 or WHO5), patient-reported outcome (PRO) or medication compliance.
- *Predictive modeling (PRE)*: technologies relying on collected data (active user inputs and/or passive sensor data) for higher-level feature extraction, modeling, and prediction of significant parameters and disease forecasting (e.g., mood forecasting in depression).
- *Intervention models (INT)*: technologies providing support for intervention, i.e., an active component that guides the patient in self- or blended care. Examples include support for cognitive behavioral therapy (CBT), psycho-education, or behavioral activation.
- *User interaction (UI)*: technologies addressing the design of user interfaces for SMI diseases. This includes, e.g., interface models for data collection, visualization, and analysis, as well as interface models for EMA, communication, psycho-education, and intervention in general.

These five topics can be viewed as a technology stack as illustrated in Figure 1. The lowest



**Figure 2.** Historical overview of the number of systems and their technology focus.

level of this stack is data collection (sensing), which can be used for clinical assessment and/or predictive modeling, as well as for intervention and guidance to the patient. In this regard, 85% of the reviewed systems were built on top of sensing technologies, followed by approximately one-third of the systems (24%–35%) that addressed prediction, user-interface design, intervention, and clinical assessment (Figure 1). Note that the user interface (UI) is a separate technological topic, which was typically transversally covered with the other topics. For example, a system like MOBERO<sup>17</sup> for children with ADHD focused on clinical assessment while also addressing the design of UI for this user group.

Fourth, Table 1 shows that the majority of the systems relied on smartphones (91%). Approximately a third (26%) of the systems included wearable technologies independently or combined with smartphones. Due to its open nature, Android was the predominant platform. More than half of the systems (55%) were US-based, 36% were EU-based, and the rest (8%) from Asia/Australia.

Figure 2 shows a historical outline of the 46 systems. Viewed from a historical perspective, it is interesting to notice that research in mobile technologies for mental health emerged early after the advent of the very first smartphones (the first iPhone appeared in 2007 and the first

## LEADER OPINION

**A**s part of this review, we interviewed nine researchers whose work has been core in establishing the research agenda on tech-supported mental healthcare and we gathered their opinion about opportunities, gaps and future direction in this field. The interviews unpacked technological issues, reproducibility, and clinical adoption as the key barriers and challenges to address in UbiComp research for mental health. Below are some of the core statements.

### Technological Issues

*"Recent tendency of mobile platforms to become progressively locked and their sensors increasingly inaccessible,*

*presents a clear technological risk for these methods to achieve their full potential"* — Venet Osmani

*"Part of the problem is also the enormous amount of heterogeneity in the data geographic, culture, age, devices, weather & climate, economic status, health, daily rhythms, etc."* — David Mohr

*"We know that we can extract some signals we need rigorous metrics, experimental protocols and benchmarks datasets to compare the algorithms."* — Mirco Musolesi

### Reproducibility

*"I think the key challenge now is reproducibility of models/ results to see if these approaches used in small scale*

*studies generalize out to a broader population.*” — Andrew Campbell

*“Computer science values novelty and innovation. But to make these tools scalable, we must have reproducibility and reliability”* — David Mohr

*It is very difficult to see how the traditional RCT approach favoured in the medical community can work in the context of new technologies. Developing new approaches that are more flexibly, but are also sufficiently robust to gain acceptance in the medical community will be an important step in realising the impact of ubicomp technologies in mental health.* — David Coyle

*“In our own work with Purple Robot and then using the StudentLife data set, we’ve been able to replicate our findings but then our models fell apart with a subsequent data collection. How do the different ways people use these devices impact that findings? Especially in some sub-populations phone use might be different (shared phones, outdated phones, etc.) so what are the boundary conditions of our findings and how do we need to address these to scale?”* — Stephen Schueller

*“The biggest impact will be common data and analytical standards that will allow for easier collaboration and replication. Also, now partnering with psychiatrists to ensure that these tools are developed to not only be*

*patient centered (or clinician centered) but rather relationship centered will be critical.*” — John Torous

## Clinical Adoption

*“I think we have generally solved the efficiency collection of data issues system-wise. I am not sure we have solved the data interpretation nor the behaviour intervention efficiency”* — Cecilia Mascolo

*“We dont really know how to use this information to the benefit of users.”* — David Mohr

*“Why do I have to go to doctor to get the technology if I can go to an electronic shop and buy it? We can look at it [data acquisition] the other way around - using consumer products and bringing the data to clinicians. Certainly, it might not be the right model for everything, but it may work, for instance, in the context of dementia and problems of older adults – their family members are primary customers.”* — Alex Mihailidis

*“How can these tools be used in clinical settings? Are they meant to replace traditional treatments and if so, do consumers want that? Will health systems accept it? If they are meant to supplement and integrate into clinical care, can clinicians understand how to use them and in what ways will they improve care (make it more efficient, make it more impactful?”* — Stephen Schueller

stable Android phone in 2009). For example, non-clinical studies of Psychlog<sup>18</sup> and StressSense<sup>19</sup> focusing on stress detection, and MoodSense<sup>20</sup> for mood sensing were published in 2010–2012. Similarly, the first clinical studies using Mobylyze! for depression tracking<sup>21</sup> and MONARCA for bipolar mood sensing and clinical assessment<sup>22,23</sup> were published a few years later, in 2011 and 2012, respectively. The number of systems and corresponding publications per year suggests two “waves” of technologies—an early wave in 2010–2012 and the second wave during 2013–2016, which was gaining more momentum.

## DISCUSSION

The review of the 46 systems above revealed that UbiComp research in mental health has opened a myriad of opportunities that hold a promise to radically change clinical practice, including:

- mitigating the gap between the demand and the provision of mental health services;

- enabling a historical shift from traditionally episodic to technology-supported continuous mental healthcare;
- providing new flexible and more personalized mental health services and interventions;
- improving accessibility to mental health services with respect to reduced costs and geographical constraints.

The last decade has brought a solid proof-of-concept for a series of technologies and we are slowly witnessing their initial clinical implementations. However, despite this significant body of research, recent systematic reviews (both technological<sup>11</sup> and clinical<sup>24</sup>) as well as our interviews with the domain experts (see the Sidebar) strongly suggest that, in practice, we are still far from the continuous technology-enabled mental healthcare model. Consistent with the literature review, the interviews highlighted the need to address technical issues, reproducibility of studies, and clinical evidence and adoption as the key challenges toward enabling the vision of

extending mental healthcare delivery beyond the traditional healthcare model.

### Technical Challenges

Advances in miniaturization and computational power of mobile and wearable devices also come with problems in the quality of sensor data, specifically data gaps, noisy readings, and a lack of precision. This stems from the fact that most of the sensors, in particular those embedded in smartphones, are not purposefully designed for behavioral modeling and for mental health applications. As pointed out by Osmani, modern smartphone manufacturers increasingly restrict access to sensors or they require a special permission, which is a trend that has been recently manifested by both iPhone and Android platforms. Although Android has been a preferable platform in the mobile sensing research as observed in our review, it may not be the case in the future. Such restrictive policies mainly aim to protect the user, however, they limit the research and understanding of potential value that the smartphones can bring in mental health. Custom sensing devices might be a viable path in the future to overcome some of these challenges, though we may need to repeat some of the previous research that relied on smartphones (as argued by Schueller) and address a potential social stigma (indicated by Osmani).

Across different systems, a number of different data categories have been collected from smartphones (e.g., accelerometer, screen ON-OFF, calls and SMS, audio signals, light, etc.) and wearables (e.g., steps, sleep quality, and duration, electrodermal activity, heart rate). A recent review<sup>11</sup> showed the correlations between such sensor-based data types and depressive symptoms. The authors reported that several studies showed consistent and statistically significant correlations. However, the evidence was limited due to the methodological issues, the lack of standardization of the collected data, mood assessment methodology, and the applied statistical methods. Therefore, the future research should resolve consistency in data collection and analysis, thus opening the door to replication studies as well as meta-analyses. In this line, Musolesi added that we are also missing rigorous metrics, experimental protocols, and benchmark

datasets to compare the algorithms. Only now after a decade of using smartphones and wearables in monitoring mental health conditions, there are pioneering efforts in establishing common agreements in defining data metrics (such as<sup>2</sup> and<sup>25</sup>) It will be increasingly important to expand such initiatives in order to validate findings across research studies conducted in different settings.

### Reproducibility

Challenges in developing predictive algorithms and understanding their reproducibility directly stem from the technical issues discussed above, yet it entails more complex challenges beyond the technical issues. Reproducibility was emphasized by five research leaders (Campbell, Mohr, Musolesi, Shueller, Torous) who raised doubts in generalizability of results achieved in small-scale studies.

Schueller pointed out one of the basic challenges in predictive modeling—the fact that different subpopulations may exhibit different patterns of smartphones use that can, in turn, impact findings. More fundamental methodological problem is related to a gap between UbiComp and clinical research, which was raised by Coyle. The challenges of combining hypothesis driven clinical research with the more exploratory approach in UbiComp was also discussed by Schueller and Mohr in.<sup>26</sup> In addition, Mohr criticized conventions in the scientific community and the fact that computer science does not encourage validation studies.

Developing predictive models and assessing their replicability highly depends on the acquisition of reliable ground-truth information in the mental health context. Unlike in many other medical domains, psychiatrists do not have objective instruments (such as an X-ray or MRI) to diagnose the patient's mental health state. Diagnosis is often based on subjective "pen and paper" based questionnaires done either by the patient or by a trained clinician. This means that diagnosis is often a difficult and time-consuming process with potential intersubjective variability across psychiatrists. This may lead to an unreliable ground-truth, which significantly impacts the identification of features and digital biomarkers, and ultimately the applicability of predictive models.

An additional challenge in developing solid and replicable models is related to a low variability of symptoms captured during the study. For instance, a number of manic episodes is typically small in comparison to the number of depressive episodes, for which discarding manic states from the prediction task sometimes becomes necessary.<sup>27</sup> A low variability in symptoms may come from the phenomenology of a specific disorder, but also from a low duration of the study and limited access to diagnosed patients.

#### Clinical Evidence and Adoption

When it comes to almost any technological innovation, acceptance, and adoption typically represent fundamental research topics. In this regard, clinical exploitation of UbiComp technologies is no different. The interviews with the domain experts and the literature review signals a “call-to-arms” to expand work on establishing the right uses. Evidently, clinical practice has not undergone radical changes, and Mohr argued that it is still not clear how information collected from mobile and wearable technologies can benefit patients. In a similar vein, Mascolo underlined that the data interpretation is a challenge, which still remains unsolved, coupled with creation of effective behavioral interventions. Mihailidis opened a parallel track to clinical use-cases—developing commercial off-the-shelf systems and tech platforms for mental health. Ultimately, what underpins the adoption of UbiComp technologies in mental health management is a solid clinical evidence which, to a certain extent, encompasses most of the discussed challenges. Recent reviews (e.g.,<sup>24</sup>) and the National Institute of Mental Health (NIMH) in the US (<https://www.nimh.nih.gov/health/topics/technology-and-the-future-of-mental-health-treatment/index.shtml>) highlight the lack of clinical evidence as the reason for which we still do not see UbiComp technologies in mental healthcare practice.

A growing number of studies that demonstrated the opportunity of using mobile and wearable technologies in mental health was not followed by a similar number of studies establishing clinical evidence of their efficacy. Expected concerns related to the lack of clinical evidence primarily come from a low sample size, sample biases, variability in the study data sample, and a

short study and deployment duration. As shown in Table 1, more than half of the systems with clinical focus included 20 participants or less, and also more than half of the studies had an evaluation period shorter than 2 months. These issues are directly related to a lack of evidence for the intervention efficacy or the reproducibility of sensing and predictive models.

In practice, conducting technological studies in clinical settings with a relatively large number of patients is by no means trivial. The literature frequently referred to the “tech-phobia” of health-care professionals who may be unfamiliar with the technology, and thus “anxious” about its use.<sup>28</sup> This may be taking a toll both on the length of Internal Review Board (IRB) and ethical reviews but also on the uptake in practical use of new technological solutions. Another important factor that impedes data collection and acquiring practical evidence is the patient’s engagement and retention. Completing questionnaires, surveys, and diaries require a continuous daily effort without immediate benefits, which might have been a contributing factor for dropping-out from the studies.

Even when the studies or mobile applications bring scientific evidence about their efficacy, there are considerable issues related to publication bias.<sup>28</sup> Whether or not a technology promises more than it delivers is commonly questioned, and NIMH also emphasizes the importance of understanding “for whom and for what” the technology is indented. The NIMH advocates for an improved understanding to what degree the technology works for what kind of patients and for what kind of conditions, which is typically not addressed in the more generic presentation of technologies.

Although experimental designs and conditions in the most of research studies sufficed for feasibility assessment, understanding clinical value still requires further research as well as a set of action points to enable the evidence of intervention efficacy and predictive models.

#### OUTLOOK

Significant research in using ubiquitous computing technologies for mental health have been conducted over the last decade and promising results—both technological and clinically—have been demonstrated. However, as discussed



above, there are still significant challenges that prevent deployment of a technology-enabled continuous mental healthcare model. What lies ahead of us and what are the important and exciting research topics for the next decade?

From a technological point of view, significant results have been achieved in mobile sensing, feature extraction, and disease prediction. In order to take this into a broader clinical practice, further research needs to tackle the heterogeneity in data and features, their link to the validated clinical measures, and the lack of reproducibility. One important technological research agenda is to further investigate dedicated “mental health” wearables. Instead of using sensors in the existing wearables and smartphones opportunistically, dedicated hardware and software platforms would embed a purposefully selected bundle of sensors tuned to a specific mental health use-case in order to deliver a consistent collection of relevant data and stable reproducible models. We might be at the dawn of expansion of medical devices for mental health that will be queuing for the US Federal Drug Administration (FDA) or related approvals. Following up the strong research focus on smartphone sensing as shown in this review, the upcoming research may focus on investigating the existing or building new types of sensing devices, such as wearable (including glasses) and implantable devices. Furthermore, it is essential to initiate standardization of features and data types, for instance following the example of the Open mHealth initiative.<sup>2</sup> Yet, this initiative currently focuses on somatic diseases, and there is a clear need to expand it to mental health. With respect to data collection, another important topic is to address the challenges of privacy-preserving data sampling, and compliance to regulatory frameworks such as the General Data Protection Regulation (GDPR) in the EU. In this regard, the ongoing research done in distributed and privacy-preserving machine learning might be exploited.

From a clinical point of view, future research undoubtedly needs to address the (lack of) clinical evidence and reproducibility. This depends on consistency in data collection, modeling, and feature extraction as well as on consistency in ground-truth labeling. Data sharing and cross-validation across multiple studies would speed up acquiring the evidence on reproducibility. As pointed out by

Schueller, only one replication study has been conducted and it yielded limited results. On a more generic level, there is a very limited evidence for the clinical efficacy of UbiComp technology for mental health. Mohr and Schueller argued that it is still unclear what the best UbiComp use-case in the clinical settings would be. Hence, future research should address the related challenges, investigate the right “clinical-socio-technical” setup for this kind of technologies, and establish its usefulness in clinical pathways—screening, prevention, treatment, care, or rehabilitation.

As such, there is a great deal of research still to be done in academic, clinical, and industrial labs, in order to reach the initial dream of technology for mental health, as set out a decade ago. So far, UbiComp has been successful in providing important technological contributions, however, when it comes to clinical usefulness “the jury is still out” (Campbell).

## ACKNOWLEDGMENT

The authors would like to thank the following researchers for their kind and insightful contributions (in alphabetical order):

- Andrew Campbell, Dartmouth College, USA.
- David Coyle, University College Dublin, Ireland
- Cecilia Mascolo, Cambridge University, UK.
- Alex Mihailidis, University of Toronto, Canada
- David Mohr, Northwestern University, USA.
- Mirco Musolesi, University College London, UK.
- Venet Osmani, Fondazione Bruno Kessler, Italy.
- Stephen Schueller, University of California, Irvine, USA.
- John Torous, Harvard University, USA.

## REFERENCES

1. D. Vigo, G. Thornicroft, and R. Atun, “Estimating the true global burden of mental illness,” *Lancet Psychiatry*, vol. 3, no. 2, pp. 171–178, 2016.
2. D. Estrin and I. Sim, “Open mHealth architecture: An engine for health care innovation,” *Science*, vol. 330, no. 6005, pp. 759–760, 2010.
3. B. Arrnrich, O. Mayora, J. Bardram, and G. Tröster, “Pervasive healthcare: Paving the way for a pervasive, user-centered and preventive healthcare model,” *Methods Inf. Med.*, vol. 49, no. 1, pp. 67–73, 2010. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/20011810>

4. T. R. Insel, "Digital phenotyping: Technology for a new science of behavior," *JAMA*, vol. 318, no. 13, pp. 1215–1216, 2017. [Online]. Available: + <http://dx.doi.org/10.1001/jama.2017.11295>
5. N. D. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, and A. T. Campbell, "A survey of mobile phone sensing," *IEEE Commun. Mag.*, vol. 48, no. 9, Sep. 2010.
6. K. K. Rachuri, M. Musolesi, C. Mascolo, P. J. Rentfrow, C. Longworth, and A. Aucinas, "EmotionSense: A mobile phones based adaptive platform for experimental social psychology research," in *Proc. 12th ACM Int. Conf. Ubiquitous Comput.*, 2010, pp. 281–290.
7. C. Wohlin, "Guidelines for snowballing in systematic literature studies and a replication in software engineering," in *Proc. 18th Int. Conf. Eval. Assessment Softw. Eng.*, 2014, Art. no. 38.
8. T. Donker, K. Petrie, J. Proudfoot, J. Clarke, M.-R. Birch, and H. Christensen, "Smartphones for smarter delivery of mental health programs: A systematic review." *J. Med. Internet Res.*, vol. 15, no. 11, Art. no. e247, Jan. 2013. [Online]. Available: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3841358&to=ol=pmcentrez&rendertype=abstract>
9. E. Dogan, C. Sander, X. Wagner, U. Hegerl, and E. Kohls, "Smartphone-based monitoring of objective and subjective data in affective disorders: Where are we and where are we going? Systematic review," *J. Med. Internet Res.*, vol. 19, no. 7, Jul 2017, Art. no. e262. [Online]. Available: <http://www.jmir.org/2017/7/e262/>
10. P. Sanches *et al.*, "HCI and affective health. Taking stock of a decade of studies and charting future research directions," in *Proc. CHI Conf. Human Factors Comput. Syst.*, ACM, 2019.
11. D. A. Rohani, M. Faurholt-Jepsen, L. V. Kessing, and J. E. Bardram, "Correlations between objective behavioral features collected from mobile and wearable devices and depressive mood symptoms in patients with affective disorders: Systematic review," *JMIR Mhealth Uhealth*, vol. 6, no. 8, Aug 2018, Art. no. e165. [Online]. Available: <http://mhealth.jmir.org/2018/8/e165/>
12. J. Seppälä *et al.*, "Mobile phone and wearable sensor-based mHealth approaches for psychiatric disorders and symptoms: Systematic review," *JMIR Mental Health*, vol. 6, no. 2, Feb. 2019, Art. no. e9819. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/30785404>
13. M. H. Aung, M. Matthews, and T. Choudhury, "Sensing behavioral symptoms of mental health and delivering personalized interventions using mobile technologies," *Depression Anxiety*, vol. 34, no. 7, pp. 603–609, 2017.
14. G. M. Harari, S. R. Müller, M. S. Aung, and P. J. Rentfrow, "Smartphone sensing methods for studying behavior in everyday life," *Current Opinion Behav. Sci.*, vol. 18, pp. 83–90, 2017.
15. World Health Organization, *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*. Geneva, Switzerland: World Health Organization, 1992.
16. J. E. Bardram, "Pervasive healthcare as a scientific discipline," *Methods Inf. Med.*, vol. 3, no. 47, pp. 129–142, 2008. [Online]. Available: <http://dx.doi.org/10.3414/ME9107>
17. T. Sonne, J. Müller, P. Marshall, C. Obel, and K. Grønbo, "Changing family practices with assistive technology: MOBERO improves morning and bedtime routines for children with ADHD," in *Proc. CHI Conf. Human Factors Comput. Syst.*, ACM, 2016, pp. 152–164.
18. A. Gaggioli *et al.*, "A mobile data collection platform for mental health research," *Pers. Ubiquitous Comput.*, vol. 17, no. 2, pp. 241–251, 2013.
19. H. Lu *et al.*, "StressSense: Detecting stress in unconstrained acoustic environments using smartphones," in *Proc. ACM Conf. Ubiquitous Comput.*, 2012, pp. 351–360.
20. R. LiKamWa, Y. Liu, N. D. Lane, and L. Zhong, "Can your smartphone infer your mood," in *Proc. PhoneSense Workshop*, 2011, pp. 1–5.
21. M. N. Burns *et al.*, "Harnessing context sensing to develop a mobile intervention for depression," *J. Med. Internet Res.*, vol. 13, no. 3, 2011.
22. J. E. Bardram, M. Frost, K. Szanto, and G. Marcu, "The MONARCA self-assessment system: A persuasive personal monitoring system for bipolar patients," in *Proc. 2nd ACM SIGHIT Int. Health Inform. Symp.*, 2012, pp. 21–30. [Online]. Available: <http://doi.acm.org/10.1145/2110363.2110370>
23. J. E. Bardram, M. Frost, K. Szanto, M. Faurholt-Jepsen, M. Vinberg, and L. V. Kessing, "Designing mobile health technology for bipolar disorder: A Field Trial of the MONARCA System," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, 2013, pp. 2627–2636. [Online]. Available: <http://doi.acm.org/10.1145/2470654.2481364>
24. J. Firth *et al.*, "The efficacy of smartphone-based mental health interventions for depressive symptoms: A meta-analysis of randomized controlled trials," *World Psychiatry*, vol. 16, no. 3, pp. 287–298, 2017.

25. A. Doryab, P. Chikarsel, X. Liu, and A. K. Dey, "Extraction of behavioral features from smartphone and wearable data," 2018, *arXiv:1812.10394*.
26. D. C. Mohr, M. Zhang, and S. M. Schueller, "Personal sensing: Understanding mental health using ubiquitous sensors and machine learning," *Annu. Rev. Clin. Psychol.*, vol. 13, pp. 23–47, 2017.
27. M. Constantinides, J. Busk, A. Matic, M. Faurholt-Jepsen, L. V. Kessing, and J. E. Bardram, "Personalized versus generic mood prediction models in bipolar disorder," in *Proc. ACM Int. Joint Conf. Int. Symp. Pervasive Ubiquitous Comput. Wearable Comput.*, 2018, pp. 1700–1707.
28. S. Lal and C. E. Adair, "E-mental health: A rapid review of the literature," *Psychiatric Serv.*, vol. 65, no. 1, pp. 24–32, 2014.

**Jakob E. Bardram** is a professor in computer science at the Technical University of Denmark [www.dtu.dk] and is the director of the Copenhagen Center for Health Technology [www.cachet.dk]. His research interests include software architecture for mobile and ubiquitous computing, mobile sensing, and human-computer interaction. He holds a Ph.D. degree from the University of Aarhus, Denmark. Contact him at jakba@dtu.dk

**Aleksandar Matic** is the head of Research Team at Telefonica Alpha—Health Moonshot, Spain. His research interests include mobile and ubiquitous computing, behavioral analysis, and pervasive health-care. He received a Ph.D. in Telecommunications from the University of Trento, Italy. Contact him at: aleksandar.matic@telefonica.com.