npg

Perspective

Harnessing Smartphone-Based Digital Phenotyping to Enhance Behavioral and Mental Health

Jukka-Pekka Onnela*,[1] and Scott L Rauch[2,3,4]

[1]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Harvard University, MA, USA; [2]Department of Psychiatry, McLean Hospital, Boston, MA, USA; [3]Department of Psychiatry, Harvard Medical School, Boston, MA, USA; [4]Department of Psychiatry and Mental Health, Partners Healthcare, Boston, MA, USA

Technology has a long track record of enabling and accelerating scientific discovery and innovation. Two prime examples are the telescope and the microscope; although these predate the industrial revolution, both can be found in today's laboratories throughout the world and they remain at the center of our scientific endeavors. In a similar way, as scientists depend on these instruments for observing the natural world, most of us have come to depend on our smartphones as tools for organizing and navigating our daily lives. Our use of these devices generates, as a byproduct, a surprisingly rich tapestry of social and behavioral fingerprints. Given that these digital fingerprints reflect the lived experiences of people in their natural environments, with granular temporal resolution, it might be possible to leverage them to develop precise and temporally dynamic disease phenotypes and markers to diagnose and treat psychiatric and other illnesses.

We have recently defined digital phenotyping as the 'moment-by-moment quantification of the individual-level human phenotype *in situ* using data from personal digital devices' (Torous *et al*, 2015). We present here an overview of smartphone-based digital phenotyping in the context of mental health (Figure 1) and outline some challenges that we foresee. We also identify specific areas within psychiatry where digital phenotyping could likely be usefully applied, and we speculate about how it might be used in the near future to improve patient outcomes and cost-efficiency by early identification of at-risk individuals. In the bigger picture, we see digital phenotyping as being part of deep phenotyping. It is also closely aligned with the goals of precision medicine, by making possible the collection of new types of phenotypic data that may be more effectively linked with genotype to identify potential connections between disease subtypes and their genetic variations (Delude, 2015).

*Correspondence: JP Onnela, Department of Biostatistics, Harvard T.H. Chan School of Public Health, Harvard University, 655 Huntington Avenue, Boston, MA, 02115, USA; Tel: +6176825781, E-mail: onnela@hsph.harvard.edu

ORIGINS OF THE TECHNOLOGY

Smartphones and other digital devices owe their existence to the development of the transistor and how they can be assembled in large numbers on small integrated circuits. First invented in 1947, the number of transistors that can be fit on a chip has approximately doubled every 2 years for the last 40 years. The heuristic that describes this exponential growth is commonly known as Moore's law, and it serves as the foundation for digital phenotyping for two reasons. First, digital sensors, enabled by the advent of micro-electro-mechanical systems, are now small and ubiquitous, which enables the collection of data from subjects in their naturalistic settings using smartphones and other personal digital devices. Second, computers now have the necessary computational capacity and sufficient amount of memory to process the sensor and other types of data using sophisticated statistical and machine learning algorithms to turn billions of bits of data into biomedical knowledge and clinical insights.

The year 2014 marked a historic moment in that for the first time there were more active mobile phone subscriptions globally than there were people on the planet (Boren, 2015). Moreover, the rate of smartphone ownership has rapidly increased over the past few years; in the United States, smartphone ownership among adults increased from 35% in 2011 to 64% in 2015 (Smith, 2015). Very few smartphones were in use until the introduction of Apple's iPhone (running the iOS operating system) in 2007, and Google's Android operating system in 2008. These two platforms have come to dominate the present smartphone market, with Android having a worldwide market share in 2015 of 83% and iOS 14%. These two platforms combined currently cover 97% of the global smartphone market, creating unprecedented opportunities to interface with large numbers of subjects. Smartphones are now used for much more than calling, texting, or basic Internet browsing. For example, 67% of smartphone owners use their phone at least occasionally for turn-by-turn navigation while driving; 62% have used their phone in the past year to look up information about a health condition; and 30% have used their smartphone to take a class or get educational content (Smith, 2015). In
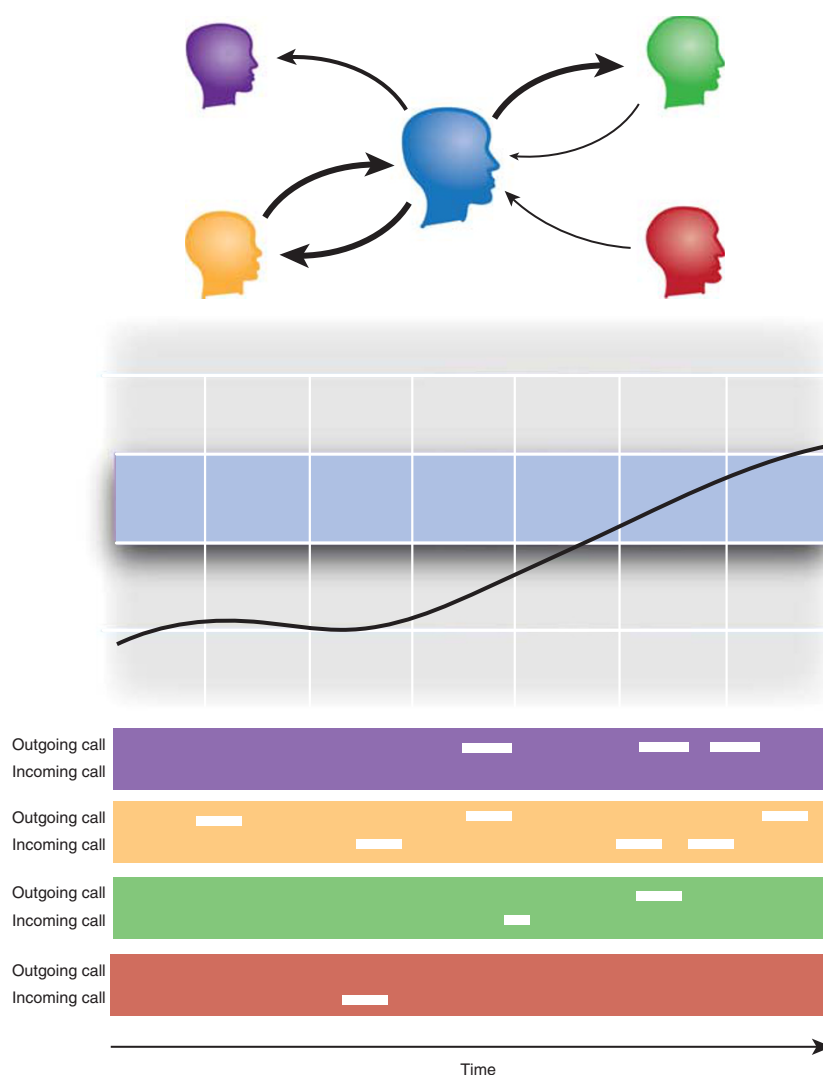
Figure 1 Schematic showing one aspect of how smartphone-based digital phenotyping may be used to learn about a subject's phone-mediated communication patterns and how those patterns might change over time in a hypothetical patient. The study subject is shown in blue in the top panel and the arrows between the subject and his social contacts represent communication flows over a period of time. Anonymized phone call logs (bottom panel) capture the underlying communication events, where each horizontal bar represents incoming and outgoing calls between the subject and one of his contacts. These data can be used to investigate call reciprocity (middle panel), a measure reflecting the balance between incoming and outgoing communication flows. A depressed subject might have reduced likelihood of returning calls, leading to low call reciprocity, but, as the patient improves, his call reciprocity might be expected to return to typical levels (blue band).

addition to scalability, the personal nature of these devices generates unique opportunities for quantifying human behavior, and the act of measurement no longer needs to be confined to clinics or research laboratories, but, instead, can be carried out in naturalistic settings–in situ–leveraging the lived experiences of subjects. Further, this can be accomplished without the incremental expense and personal burden associated with additional specialized equipment.

MAKING SENSE OF SMARTPHONE DATA

Smartphone-based digital phenotyping encompasses the collection of a range of different behavioral data, including, but not limited to, spatial trajectories (via GPS), physical mobility patterns (via accelerometer), social networks and social dynamics (via call and text logs and Bluetooth), and

voice samples (via microphone) (Torous et al, 2015). Given that smartphone sensors and phone usage patterns generate complex longitudinal multivariate data, to better understand the underpinning of digital phenotyping, it is useful to consider what types of raw data are collected. We have previously divided the data collected by a digital phenotyping smartphone application into two categories: active data (such as surveys), requiring active participation from the subject to be generated, and passive data (such as GPS traces), generated without any participation or action from the subject.

Active and passive data from smartphones, when coupled with appropriate analytical methods, could shed light on many questions of both scientific and clinical interest. GPS data, for example, could be used to learn how a depressed patient divides her time between distinct locations, such as home and work, and how both the number of locations

and time spent at these locations change in time. Phone communication logs could convey information about the size and reciprocity of a person's social networks, and these social markers could indicate cycling between depression and mania for bipolar patients. Speech samples recorded using the phone's microphone could be used to detect vocal markers of mood or they might have prognostic value for neurological disorders characterized by speech impairment. Finally, accelerometer data could be used broadly to quantify physical mobility patterns and metabolic expenditure in a range of patients, and more targeted uses of these data could focus on quantifying tremors in patients suffering from involuntary muscle contractions. Passively collected social and behavioral data might also be less susceptible to the complexities introduced by potential linguistic and cultural barriers than more traditional surveys.

The data generated by smartphone sensors and phone use patterns appear ideal for capturing various social and behavioral dimensions of psychiatric diseases, but at first the variety and types of data generated appear daunting. For example, the accelerometer can generate hundreds of thousands of observations per day, GPS can be sampled thousands of times a day, and smartphone microphones can record high-quality audio data at the rate of a compact disc. This complexity of data makes it clear why smartphone-based digital phenotyping must rely upon raw data. While it is possible to summarize such high-dimensional data using a small set of summary statistics, the choice of these summaries needs to be driven by specific scientific questions and statistical considerations. Put differently, while complex data can be summarized in different ways, not all summaries are equally useful, and one-size-fits-all summaries are unlikely to leverage the full potential of the data.

As digital phenotyping matures, it seems likely that the intellectual challenge will shift from data collection to data analysis and modeling. This observation parallels some recent developments in DNA sequencing, where the field of statistical genetics emerged to develop and apply statistical methodology to draw inferences from genetic data. There is also an analogous history for psychiatric brain imaging, whereby rich spatiotemporal data sets were progressively integrated with clinical data, and spawned a whole new field around experimental design, data analytics, and interpretation of results. One can envision a similar trajectory for digital phenotyping, where traditional statistical tools might be combined with machine learning to translate smartphone sensor and usage data into biomedical and clinical insights. Most existing approaches used to quantify measurements from mobile phones (not just smartphones), such as those used to estimate mobility, appear to ignore key statistical considerations like dealing with missing data and quantifying uncertainty of the estimates. We expect the development of statistical learning techniques that are specific to the domain of digital phenotyping, and driven by specific scientific questions in the area, to be a major intellectual undertaking in the near future. Finally, more precise disease phenotypes from digital phenotyping might be successfully combined with DNA and RNA sequencing and with epigenetics. This could contribute to the development of a classification of psychopathologies based on dimensions of observable behavior that integrates information from genetics, making it compatible with the research domain criteria framework of the National Institute of Mental Health in the United States.

RELATED APPROACHES

Psychiatry has a history of quantifying human behavior *in situ*, and previously this has been achieved by requiring subjects to carry dedicated devices for the duration of the study. Ecological momentary assessment (EMA) refers to a collection of methods used in behavioral medicine research for participants to report on symptoms and behaviors close in time to experience and in the participant's natural environment (Shiffman *et al*, 2008). It is related to the experience sampling method (ESM) (Csikszentmihalyi and Larson, 1987).

There are some important distinctions between digital phenotyping and the EMA/ESM approach. (Note that digital phenotyping is distinct from the 'digital phenotype' that was introduced recently (Jain *et al*, 2015).) First, digital phenotyping can rely on a range of passive behavioral data, without any participation from the subject, to collect a range of relevant social and behavioral markers, such as spatial trajectories, physical mobility patterns, and social dynamics. In contrast, EMA/ESM relies on self-reported accounts of behavior. Second, if active participation from the user is requested, these inputs can include surveys but do not have to be limited to them; it is easy, for example, to collect other types of active data, such as audio samples. Third, digital phenotyping uses smartphones instead of specialized research devices, which makes it substantially more scalable than traditional implementations of EMA, such as those using personal digital assistants. Neither traditional paper-based surveys nor clinical interviews were designed to be employed multiple times a day, and paper as a medium for delivering surveys is not easily customizable. Smartphones make it possible to implement adaptive, high-frequency surveys that reduce respondent burden while allowing for long-term follow-up to learn about the variability of symptoms at different time scales. Digital phenotyping can also go beyond smartphone-based implementation of EMA; for example, it is possible to have passively collected data trigger context-sensitive micro-surveys, which could be tied to specific events (eg, phone calls via communication logs) or locations (via GPS). Fourth, digital phenotyping is less focused on bringing surveys to subjects but instead attempts to capture, with minimal interference, different aspects of the ways in which the subjects interact with the surrounding world. As smartphone technology evolves, it will likely be possible to capture more and more details about these interactions. This is also why we prefer the term 'phenotyping' to 'phenotype'; the latter suggests a static construct, like the genotype, whereas the former emphasizes its dynamic nature.

Beyond EMA/ESM, mobile health (or mHealth) oriented smartphone applications provide platforms for tracking almost anything related to health from nutrition to exercise to sleep, most of them relying on user-reported data, although an increasing number of applications now leverage passive data towards specific goals, such as counting steps. Very recently mHealth applications have been used in intervention studies to increase the user's engagement in

activities that reduce their symptoms, improve adherence to medication, or support self-management of their illness (Ben-Zeev *et al*, 2015). Outside of health, mobile phone sensing is an area of computer science that uses mobile devices, more recently smartphones, to provide microscopic and macroscopic views of cities, communities, and individuals (Lane *et al*, 2010). Mobile sensing makes heavy use of sensor data, but does not encompass clinical or genetic data and tends to assume more algorithmic as opposed to statistical approaches to data analysis and interpretation of results.

THE CHALLENGE OF PSYCHIATRY: GAPS AND POTENTIAL SOLUTIONS

Twenty-first-century psychiatry presents a ripe landscape fraught with challenges for which digital phenotyping may offer novel solutions. Psychiatric disorders are common with a lifetime prevalence of over 30%, such that virtually every family in the United States is affected. These conditions, being among the top causes of disability worldwide, confer high levels of morbidity and mortality, are principal drivers of high suicide rates, and are associated with shortened life expectancy, such as in the chronically mentally ill. While there is no question that progress in this domain should be at the top of the public health agenda, historically, in part related to the stigma of psychiatric illness, mental health care has been poorly reimbursed under a fee for service model of payment. Consequently, the field has suffered from inadequate service capacity (eg, too few psychiatric beds and limited access to outpatient care) as well as fragmentation of services. Over the past decade there has been a notable shift in US health care towards new payment models emphasizing cost-efficiency and enhanced care coordination. In this context, there is a growing appreciation for the critical nature of addressing psychiatric conditions to enhance overall health outcomes and total medical spending (Goodell *et al*, 2011). The time is now especially right for investment in novel solutions in psychiatry that could deliver higher quality through enhanced outcomes, improvements in cost-efficiency, or both.

Many psychiatric disorders are characterized by a chronic or recurrent course—prime examples include major depression or bipolar illness, alcohol or other substance-use disorders, and schizophrenia or other psychotic disorders. Conventional models of care have typically awaited major psychiatric crises, leading to an emergency room visit, followed by a brief inpatient hospitalization, and then discharge to outpatient care (when available), with staggeringly high readmission rates. Emerging models of care include screening efforts to enable early detection and system development to achieve a more seamless continuum of services.

An enduring paramount challenge in the field is to determine which ambulatory patients require heightened attention and intervention in advance of psychiatric relapse in order to avert the disruption, cost, and potential tragedy associated with repeat crises. It has traditionally been difficult to identify which patient with a history of risk of major mood disorder might become depressed or manic again; which individual suffering from a substance-use disorder is at heightened risk to use again; which patient with chronic psychosis is about to decompensate; or which patient of any kind may be at heightened risk for suicide or violence.

Digital phenotyping could enable such predictions and alert the patients themselves, a designated loved one, or a professional caregiver. Such predictive identification could ultimately enable timely and effective intervention.

THE PLAUSIBILITY OF DIGITAL PHENOTYPING AS A PREDICTOR OF PSYCHIATRIC COURSE

Assessment of the sensitivity, specificity, or predictive value of digital phenotyping as an approach to predict and identify those at risk of relapse is an empirical matter. It would however be a long and arduous path to develop tests to determine how best to enhance patient outcomes by timely intervention. Is it plausible? Interestingly, the data generated by this method represent a striking parallel to the mental status exam, with arguably superior objective indices of speech, motor activity, mood, affect, thought, and cognition. While routine outpatient follow-up may entail a weekly or monthly face-to-face evaluation (perhaps even less frequently for those deemed stable), this technology could capture these indices continuously and monitor for a subtle social or behavioral red flag. For example, decreased communication, motor activity, or altered aspects of speech with a particular change in sleep times could be the harbinger of a depressive relapse. Similarly, characteristic travel and communication patterns associated with places and people with heightened risk of substance abuse could presage heightened risk for loss of sobriety, or stereotypic communication patterns and greater social isolation could signal an impending psychotic decompensation.

Examples may best serve to illustrate the most promising areas for initial research efforts. Upon discharge from inpatient care, patients with psychotic disorders often have readmission rates exceeding 20% within 30 days (Heslin and Weiss, 2012). Likewise, those discharged from substance-use rehabilitation experience high rates of relapse. Study of these populations has great potential for identifying potential predictors of relapse or readmission. Patients receiving maintenance electroconvulsive therapy are another group where an opportunity may exist to better guide optimal frequency of treatment.

Finally, the cost-efficiency and scalability of digital phenotyping could enable identification of both individuals and specific populations at heightened risk. Retrospective analysis after sentinel events could provide a key to prevention for an individual (for relatively frequent events) or for a population (for relatively infrequent events). For example, this could apply to better predicting panic episodes in an individual with panic disorder or flashbacks in an individual with post-traumatic stress disorder; individuals at rising risk of suicide among an entire population of students; or heightened risk for acts of violence across an entire community.

PROMOTING MENTAL HEALTH *VS* COMBATING MENTAL ILLNESS

There are several convergent movements that together represent a paradigm shift from the health-care focus on

combating disease toward proactively promoting health. From positive psychology to the fitness movement, as well as a growing focus on fostering resilience to enable optimal responses to stress and trauma, it is intriguing to consider how digital phenotyping could be deployed in novel ways to accomplish enhanced mental health. If digital phenotyping is able to point to indices associated with preferred mental states and enhanced behavioral outcomes, whether at the individual or population levels, it could be plausible to provide real-time feedback to individuals, enabling them to sculpt their behavior toward enhanced mental health. In many respects this resembles the emergence of Fitbit and similar technologies to catalyze those behaviors that drive individuals toward enhanced physical health. Imagine, for instance, a synthesis of digital phenotyping indices quantifying an individual's level of social engagement and steering him or her toward optimizing these based on their own genotype, profile, and patient history.

POTENTIAL RISKS AND PITFALLS

Patient and participant privacy is always of utmost importance in clinical and research settings. Given the large quantities of different types of data that may be collected using smartphone-based digital phenotyping, it is important to give this topic serious consideration. Although perhaps an obvious point, it is worth emphasizing that a patient or participant has to give his or her consent to be enrolled in a research study or clinical setting that involves digital phenotyping, and only then can one proceed with installing and activating the app on the phone. How the collected data are handled depends entirely on the specific app and the broader platform in which it is embedded. For example, the Beiwe platform developed by one of the authors (JPO) hashes all direct identifiers, ie, it replaces phone numbers in phone call logs with surrogate keys, and additionally encrypts all data while buffered on the phone, in transit, and stored on the server (Torous et al, 2015). In this case the encryption is asymmetric, meaning that the phone cannot read its own data once encrypted. The app itself is also password protected. As health-care-related information technology security risks, solutions, and regulatory standards are all rapidly evolving, work in this space will need to be nimble and vigilant to remain acceptably safe and compliant.

Passive data also raise some special considerations for privacy, patient safety, and data analysis. For example, giving subjects feedback based on their passive data on the extent of their recent social engagement could potentially exacerbate certain types of illnesses. This feedback could additionally contaminate the data analysis component of the study by making it hard or impossible to identify whether changes in patient symptoms or behavior resulted from passive data feedback, the intervention or treatment under investigation, or both. Moreover, the very concept of having one's communications or actions tracked electronically could exacerbate symptoms of anxiety, obsessions, or paranoia. While such patients may have the insight to decline participation, some may not. Further, as such symptoms can be temporally dynamic, it is foreseeable that patients could initially agree to participate and then withdraw consent and discontinue participation precisely because of

symptom exacerbation. This raises the intriguing possibility that withdrawal of participation once initiated could be its own potent predictor or signal of disease worsening, whether or not caused by the application. In general, while virtually any inquiry into psychiatric symptoms can theoretically worsen a patient's psychiatric state, such probes, and especially unobtrusive ones, are typically viewed as low risk, but determining that is ultimately an empirical matter. In the case of digital phenotyping, efforts should be made to consider what these risks might be at the outset of any investigation, and studies should include measures to try and ascertain whether the participation in active digital phenotyping itself exacerbates any of the conditions under study.

Any technological development risks the creation of a digital divide between those with access to technology and those who do not interact or interface with technology. According to a recent report, 7% of US adults have no broadband access at home and have few options for getting online other than their smartphone (Smith, 2015). Although this evinces the existence of a digital divide in the country, it suggests that smartphones are part of the solution rather than part of the problem. The numbers are even more striking globally. While in 2015 there were 2.6 billion smartphone subscriptions in the world, the number of smartphone users is predicted to be 6.1 billion by 2020, overtaking basic fixed-phone subscriptions. These numbers highlight the greatest strength of smartphone-based digital phenotyping, its potentially massive scalability.

CONCLUSION

The persistent difficulty of accurately and reliably quantifying disease phenotypes, perhaps especially in psychiatry, represents a long-standing barrier to progress in both clinical and research missions. The heterogeneity of disease phenotypes has made it difficult to understand disease mechanisms and their underlying genetic associations, thereby limiting success at predicting clinical course or developing effective personalized treatment strategies. Smartphone-based digital phenotyping has the potential to offer psychiatry a wealth of data on disease phenotypes that could provide new leverage in each of these domains. While the scope and scale of the data permit novel approaches to health, the data alone are not enough. As the developments in fields such as genetics and brain imaging demonstrate, these vistas of big data will call for the creation of new methods of data analysis; we believe this to be one of the most immediate intellectual challenges to the approach.

Smartphone-based digital phenotyping, as part of deep phenotyping, appears very promising for making progress in genomic approaches to precision medicine. By contributing to the richness of disease phenotypes, it has the potential to advance medicine more broadly, and in particular with regard to complex diseases. These are incredibly exciting times at the nexus of psychiatry, technology, and the quantitative sciences. We anticipate there to be an escalating investment of intellectual and material resources in this area based on the potential for fundamental disruption in a sector of health care where such is sorely needed. The cost-efficient and inherently scalable nature of smartphone-based digital phenotyping suggests that it could make a significant contribution to this field. The extent to which digital

phenotyping may fulfill its potential as a tool for advancing research and care in psychiatry, as well as for enhancing mental health, is an empirical question and one that, we believe, is well worth pursuing.

FUNDING AND DISCLOSURE

REFERENCES

Ben-Zeev D, Schueller SM, Begale M, Duffecy J, Kane JM, Mohr DC (2015). Strategies for mHealth research: lessons from 3 mobile intervention studies. *Adm Policy Ment Health* **42**: 157–167.

Boren ZD (2015). Active mobile phones outnumber humans for the first time. International Business Times. Available at: http://www.ibtimes.co.uk/there-are-more-gadgets-there-are-people-world-1468947.

Csikszentmihalyi M, Larson R (1987). Validity and reliability of the experience-sampling method. *J Nerv Ment Dis* **175**: 526–536.

Delude CM (2015). The details of disease. *Nature* **527**: S14–S15.

Goodell S, Druss BG, Walker ER (2011). *Mental Disorders and Medical Comorbidity*. Robert Wood Johnson Foundation: Princeton, NJ, USA. Available at: http://www.integration.samhsa.gov/workforce/mental_disorders_and_medical_comorbidity.pdf.

Heslin KC, Weiss AJ (2012). Hospital Readmissions Involving Psychiatric Disorders, 2012. Statistical Brief # 189. Agency for Healthcare Research and Quality. http://www.hcup-us.ahrq.gov/reports/statbriefs/sb189-Hospital-Readmissions-Psychiatric-Disorders-2012.pdf.

Jain SH, Powers BW, Hawkins JB, Brownstein JS (2015). The digital phenotype. *Nat Biotechnol* **33**: 462–463.

Lane ND, Miluzzo E, Lu H, Peebles D, Choudhury T, Campbell AT (2010). A survey of mobile phone sensing. *IEEE Commun Mag* **48**: 140–150.

Shiffman S, Stone AA, Hufford MR (2008). Ecological momentary assessment. *Annu Rev Clin Psychol* **4**: 1–32.

Smith A (2015). *U.S. Smartphone Use in 2015*. Pew Research Center: Washington, DC, USA. Available at: http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/.

Torous J, Kiang M, Lorme J, Onnela JP (2016). New tools for new research in psychiatry: A scalable and customizable platform to empower data driven smartphone research. *JMIR Mental Health* (in press).