Digital phenotyping for assessment and prediction of mental health outcomes: a scoping review protocol

Pier Spinazze,1,2 Yuri Rykov,1 Alex Bottle,3 Josip Car1,4

ABSTRACT

Introduction Rapid advancements in technology and the ubiquity of personal mobile digital devices have brought forth innovative methods of acquiring healthcare data. Smartphones can capture vast amounts of data both passively through inbuilt sensors or connected devices and actively via user engagement. This scoping review aims to evaluate evidence to date on the use of passive digital sensing/phenotyping in assessment and prediction of mental health.

Methods and analysis The methodological framework proposed by Arksey and O’Malley will be used to conduct the review following the five-step process. A three-step search strategy will be used: (1) Initial limited search of online databases namely, MEDLINE for literature on digital phenotyping or sensing for key terms; (2) Comprehensive literature search using all identified keywords, across all relevant electronic databases: IEEE Xplore, MEDLINE, the Cochrane Database of Systematic Reviews, PubMed, the ACM Digital Library and Web of Science Core Collection (Science Citation Index Expanded and Social Sciences Citation Index), Scopus and (3) Snowballing approach using the reference and citing lists of all identified key conceptual papers and primary studies. Data will be charted and sorted using a thematic analysis approach.

Ethics and dissemination The findings from this systematic scoping review will be reported at scientific meetings and published in a peer-reviewed journal.

INTRODUCTION

The collection of a new generation of health data is becoming ubiquitous, seamless, unobtrusive and continuous—no longer confined to clinics, laboratories or specialised medical instrumentation. All digital devices that humans interact with or are in their environment, including but not limited to the internet, computers, wearables and the Internet of Things (IoT), collect or generate data about us that can provide insights into our behaviour and health.

This process of personal digital data capture or digital phenotyping was defined as the ‘moment-by-moment quantification of the individual-level human phenotype in situ using data from digital devices’ by Torous et al., 2016.1 This is an extremely broad definition, the range of which has not been clearly identified. For the purposes of this review, we define digital phenotyping as the process of inferring individual behaviour from digital data generated through human interaction with electronic devices, including both physical hardware and software. This data can be both: (1) active, wherein the individual is required to perform a task or act to capture this data, for example, complete a survey or (2) passive, wherein there is no action or requirement of the individual outside that of normal daily activity or behaviour. This review focuses on passive data sources from mobile digital devices including sensors (from smartphones, wearables, medical or experimental devices) and electronic activity (including social media, device activity, e-stores, etc).

Smartphones

Mobile devices, including mobile phones and tablets, have become ubiquitous accessories, with now more of these devices on the planet than people.2 Considering the sheer amount of time people spend on their mobile phone, the data regarding their behaviour and health status is highly insightful and easily retrievable. An online survey study of 18–44 year olds revealed that 80% of the participants checked their phone first thing in the morning and 79% of respondents had their phone on or near them except for approximately 2 hours...
of their waking day. The extensive sensor capabilities of these devices, including tracking movement, location, light, proximity to other devices, sound, heart rate and facial recognition, allow us to amass insightful data on an individual basis.

Wearables

Beyond smartphone data, also considering wearables, such as fitness or sleep trackers, which may be acquiring uninterrupted user data as well as the IoT (IoT—wireless connectivity of all electronic devices) the volume of data which can be collected is vast. The number of wearables is expected to jump from an estimate of 325 million connected devices in 2016 to over 830 million in 2020.

Digital phenotyping has received increasing attention due to the rapid advances and potential this technology offers to healthcare. An increasing number of studies are seeking to elucidate what all this data reveals about a person’s health, behaviour and whether it can predict or measure changes in outcomes. To date, mental health has been an area of particular interest for digital phenotyping due to the very subjective nature of diagnostics within this field and the strong correlation with changes in behaviour detectable by mobile digital devices. However, digital phenotyping might be used for risk prediction of other health conditions, such as stress, addiction, wellbeing, obesity, locomotor dysfunction. Digital phenotyping holds potential value for public health, if it could scale up population health screening and improve early detection of diseases.

STUDY RATIONALE

The ubiquity of mobile digital devices provides greater penetration of the general population and provides for a focused area on which to assess individual as well as combinations of sensor data as predictive factors for health disorders. To date, there have been several published reviews in this field of research Guntuku et al; Dogan et al; Rohani et al; Reinertsen and Clifford; Cornet and Holden. However, these reviews have a number of limitations: existing reviews either focus on digital traces (see box 1) in online social networks or on smartphone and wearables sensor data. However, for better prediction power, digital phenotyping should harness digital traces from multiple different sources, and we intend to include all sources (both sensor data and electronic activity) in our review.

With the exception of Rohani et al, reviews are mostly descriptive and do not attempt to determine which digital features are the best predictors for particular health outcomes. By expanding the tech scope to include not only sensor data but electronic activity (see), we aim to add new variables for comparison.

Current reviews do not develop or apply existing frameworks of digital phenotyping (ie, frameworks of relationships between behaviours, digital traces/ sensors and health conditions, eg, suggested by Mohr et al or Garcia-Ceja et al) to highlight gaps, identify potential digital markers and generate new hypotheses. Huckvale, K. et al provide a general overview and mapping of digital phenotyping to clinical application, that is, prevention, screening, monitoring and treatment but do not attempt to map associations between digital markers and mental health outcomes. Fracarro, P. et al map the associations between digital markers and mental health outcomes; however, this review was limited to geolocation data and specific serious mental illness (ie, bipolar disorder and schizophrenia). By applying an analytical framework, we will comprehensively map the existing associations between digital features and mental disorders and try to reveal unexplored potential digital markers.

The proposed scoping review aims to overcome the above limitations, to further map this research field and establish potential digital health markers and methods for health prediction through digital phenotyping.

AIMS AND OBJECTIVES

The aim of the proposed scoping review is twofold: to summarise and map results of digital phenotyping studies on digital markers in mental health, and suggest the potential digital markers and methods for health prediction.

Digital features based on hardware (eg, sensors) and software (eg, online social networks) data sources will be assessed for effect, significance, limitations and potential areas of development.

The objectives of the review are the following:

### Box 1 Definitions

| Mental health disorder: A clinically diagnosed mental health disorder as defined by the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) criteria, for example, depression, bipolar, anxiety disorders. |
| Mobile digital device: Any electronic device that can be easily carried by an individual and is not dependent on physical connections or close vicinity to other fixed/immobile equipment in order to function, for example, smartphones, fitness trackers. |
| Sensor data: Sensor data are the output of a digital device that detects and responds to interactions with the physical environment, for example, accelerometers detect changes in gravitational acceleration of a device to determine acceleration, tilt and vibration. |
| Electronic activity: All individual activity recorded through hardware and software use, including device usage data, social media activity, browser history, cookies and online shopping. |
| Feature/variable: Feature is a quantifiable measure of some aspect of human behaviour potentially correlated with a symptom or health condition, for example, home stay—percentage of time a person spends at home relative to other locations. |
| Digital marker: Digital marker is a sustainable and interpretable statistical association between features and health outcomes or rather, digital risk factor, for example, increased home stay may indicate loss of interest in activities or decreased activity and is associated with depression. |
To identify which mental health disorders (or their symptoms) have been predicted or measured by digital phenotyping.

To identify which digital markers of human behaviour and physiology are currently detectable through personal digital devices and online activity.

To identify which digital markers have a stable statistically significant and reproducible relationship with specific mental health conditions (or their symptoms).

To group digital markers in categories that represent more high-level human behaviours and practices.

To suggest new potential digital markers.

**METHODS**

We considered several systematic review study designs, but chose to undertake a scoping review of the literature on digital phenotyping as a means to evaluate the field. This prevents us from narrowing the parameters to prespecified research trials or quality indicators as in a systematic review or meta-analysis. The scoping review methodology is a rigorous approach to mapping a research area to establish the nature, extent, range and to summarise key findings of research. It will also allow us to clarify working definitions on this topic. As this is a relatively novel and developing area of research, a scoping review appears to be the most appropriate method of assessment. Following this review, further research including a systematic review may be warranted.

We followed the Joanna Briggs Institute (JBI) recommendations and will be using the following methodological framework proposed by Arksey and O’Malley to conduct our scoping review. The framework includes the following five steps to conduct a scoping review:16:

1. Identifying the research question.
2. Identifying relevant studies.
4. Charting the data.
5. Collating, summarising and reporting the results.

This form of review is intended to be a precursor for potential further work, as on initial analysis it is unclear if a more sophisticated review method is warranted. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews guidelines will be followed for the study write up.

**Stage 1: identifying the research question**

It is important to define the research question from the outset to provide a purpose and approach to the literature search. However, this should be considered an iterative process: as the familiarity with published research increases, the question may be reframed accordingly. The primary and secondary research questions are defined in Table 1. Since we are interested in the potential population, public health value of digital phenotyping, we will limit the scope of this review to technology and data sources from personal non-medical consumer devices, widely used social web services, and passive modes of data collection, which do not require any specific action or intervention from the participants.

**Stage 2: identifying relevant studies**

A three-step search strategy will be used according to JBI recommendations. The first step is an initial limited search of online databases namely, MEDLINE for literature on digital phenotyping or sensing. A sample of relevant articles was analysed for terms used in the title, abstract and indexed terms. A second search using all these identified keywords will then be undertaken across all relevant electronic databases: IEEE Xplore, the ACM Digital Library, MEDLINE (see search strategy online supplementary appendix A), the Cochrane Database of Systematic Reviews, PubMed, Web of Science Core Collection (Science Citation Index Expanded and Social Sciences Citation Index) and Scopus. The search strategy will be modified for each database and further iterated as we explore the research question, with changes captured in the review process. Third, a snowballing approach will be taken. A list of top reviews, conceptual papers and primary studies will be created and the reference and citing lists of all identified articles will be searched for additional studies. Authors of primary studies or reviews may be contacted for further information, if deemed necessary. We will use the following inclusion criteria as per Population, Intervention, Comparator, Outcomes, Study type (PICOS) framework (table 2). However, the PICOS framework should be adjusted for observational (non-interventional) types of studies, because the use of technology is not aimed to change behaviour or to induce an effect in participants. Hence, ‘Intervention’ and ‘Comparator’ elements contain description of...
Review inclusion criteria

<table>
<thead>
<tr>
<th>Population</th>
<th>General population, human users of digital technologies (mobile digital devices including smartphones and wearables) and personal electronic activity (eg, social media activity, device usage, etc). Mental health conditions listed in DSM-5 criteria.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention (predictor)</td>
<td>Use of electronic technology through which personal behaviour can be captured and measured (eg, data collected via mobile digital devices or electronic activity for example, social media usage, etc) as a source of mental health predictors.</td>
</tr>
<tr>
<td>Comparator (outcome)</td>
<td>Use established methods of mental health diagnosis (defining health status) and assessment of mental health symptomatology as a ground truth comparator with digital phenotyping, for example, clinician assessment, medical records, screening tools and tests, patient diaries, self-reported surveys.</td>
</tr>
<tr>
<td>Outcomes</td>
<td>The accuracy of prediction or strength of statistical relationship between digital features and mental health disorders (or their symptoms).</td>
</tr>
<tr>
<td>Study type</td>
<td>Observational studies (cohort/longitudinal, case-control, cross-sectional). Only publications with abstracts in English will be included. Due to technological advancement and advent of smartphones and sensors we will only review publications from 2009.</td>
</tr>
</tbody>
</table>

Exclusion criteria

- Studies using only outcome data (eg, ecological momentary assessment (EMA), digital diaries/journals, etc) without passive electronic device data as predictors.
- Studies using only digital phenotyping data without relation to ‘ground truth’ data (validated risk assessments or diagnostic data) on health conditions (clinical assessment, screening tests, medical records, etc).
- Studies using sensor data based on non-mobile devices (eg, systems of wearable sensors for fine-grained gait analysis, clinical lab equipment, etc).
- Intervention studies (eg, cognitive–behavioural therapy interventions).
- Non-quantitative studies, for example, purely qualitative studies, or mixed-methods studies wherein the quantitative portion does not match the inclusion criteria.
- English version of the paper is unavailable.
- Studies with nonhuman participants.
- Studies with health conditions unrelated to mental health.
- Studies published before 2008.
- The following publication types: trial protocols, systematic reviews, in vitro or laboratory research, letters or opinions, conference abstracts and posters.

There are limitations to not including grey literature, randomised controlled trials (RCT) and systematic reviews, as these can potentially hold useful information however do not fit appropriately with the objective of the review, which is focused on observational studies and those that have undergone peer review. We may include these if the current search criteria yield a low number of relevant studies. Alternatively, if the search includes too many irrelevant publications, we will review the categories and combinations of terms independently to reveal where the majority of these publications come from and revise accordingly.

The initial literature search strategy used for MEDLINE can be found in online supplementary appendix A, including the free-text used to perform the search by combining both digital and health-related terms.

Stage 3: study selection

Screening of the studies will be performed by two reviewers and at two levels. Table 2 outlines the inclusion criteria that will be used by the reviewers to determine which studies will be included. The citation management software program EndNote X8.2 (Clarivate Analytics, USA) will be used to manage records and data and to remove duplicates. Initial screening will involve the title and abstracts where available, with articles categorised into include, exclude and unsure. Predefined relevance criteria will be used to determine the relevance of studies. The second round of screening will involve full-text reading of the studies identified in the previous round to ensure relevance and inclusion and reach consensus between the reviewers. Any conflicts or discrepancies will be resolved by discussing with a senior researcher.

Stage 4: charting the data (data extraction)

A standardised data collection template created by the research team with predefined criteria will be used for data extraction. All relevant studies will be reviewed independently by the two researchers (authors PS and YR) and data extracted according to preliminary fields indicated in Table 3. Broadly, these will be grouped by study details including type, year, population and digital phenotyping or measured digital metrics and validation criteria. The form will be piloted initially on a few studies to ensure that the fields are broadly applicable and assess whether there should be any additional fields included. Each researcher’s independent data abstraction will be compared and discrepancies discussed and resolved if not internally, by a third party. Additional fields may be included if during the process the researchers agree that they are relevant and should be included.

Stage 5: collating, summarising and reporting the results

Basic numerical analysis of extracted data will allow for summaries of relevant study characteristics, such as mental health disorders, technical sources of data, digital features, methodology and sample sizes. Predictive characteristics of digital features extracted from studies will
Table 3  Data fields and explanations

<table>
<thead>
<tr>
<th>Article details</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Year</strong></td>
<td>Year of publication</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study details</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study design</strong></td>
<td>Cohort, case–control, cross-sectional</td>
</tr>
<tr>
<td><strong>Sample size</strong></td>
<td>No of participants included in the study</td>
</tr>
<tr>
<td><strong>Study population</strong></td>
<td>Study population details</td>
</tr>
<tr>
<td><strong>Data source</strong></td>
<td>Digital device or online platform used to acquire participant data</td>
</tr>
<tr>
<td><strong>Study duration</strong></td>
<td>No. of days/weeks</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Digital phenotyping/sensing</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Application/platform</strong></td>
<td>Name of mobile application or platform if used</td>
</tr>
<tr>
<td><strong>Sensors</strong></td>
<td>Types of sensors used to acquire data for example, accelerometer, light sensor</td>
</tr>
<tr>
<td><strong>Features</strong></td>
<td>Features used for prediction, for example, distance travelled</td>
</tr>
<tr>
<td><strong>Type of behaviour</strong></td>
<td>What behaviour each feature elucidates for example, movement patterns, sociability, communication, device use</td>
</tr>
<tr>
<td><strong>Behavioural marker /risk factor</strong></td>
<td>What each feature indicates in terms of health indicators or risks (poor sleep, lack of movement, unstable circadian rhythms, reduction of social contacts)</td>
</tr>
<tr>
<td><strong>Methodology</strong></td>
<td>What statistical method was used to determine significance, for example, correlation, regression, classification trees, time series analysis</td>
</tr>
<tr>
<td><strong>Effect size, accuracy, measure of significance</strong></td>
<td>Standardised effect size, unstandardised effect, specificity/sensitivity (recall/precision), Area Under the Curve (AUC), R², p value</td>
</tr>
<tr>
<td><strong>Mental health disorder</strong></td>
<td>What health condition is the study attempting to measure in terms of detection and prediction?</td>
</tr>
<tr>
<td><strong>Source of validation</strong></td>
<td>What source of ‘ground truth’ data is used to evaluate the features, for example, clinician assessment, medical records, screening tests, patient diaries, self-reported scales</td>
</tr>
</tbody>
</table>

be reported in the form of summary tables mapping the associations between digital features and mental disorders. Then we will inductively aggregate identified digital features into categories that represent interpretable types of human behaviour and practices, in a way similar to the one suggested by Mohr et al (figure 1). This abstraction will provide more high-level vision on relationships between digital phenotyping and mental health disorders,

Figure 1  Analytical framework from Mohr et al 12 (GPS: Global Positioning System; SMS: short message service)
and will help to identify gaps requiring further research. In particular, this framework will highlight behavioural categories with higher value for prediction of particular health conditions and will advise what other digital indicators describing the same behavioural category can be derived from the data.

Included studies will not be assessed for quality nor methodology specifically as this is outside the scope of this type of review. However, statistical methods used for evaluation of association strength or prediction of health condition will be recorded as well as accuracy measures (sensitivity/recall, precision, specify, F-score, etc).

PATIENT AND PUBLIC INVOLVEMENT
No patients will be involved in this study.

ETHICS AND DISSEMINATION
This review contributes to the advancement of knowledge within the field of digital phenotyping which is still fairly new and experimental. To our knowledge, this scoping review is the first of its kind to summarise digital phenotyping uses in relation to a broader health scope. The findings of this scoping review will provide a comprehensive description of potential digital health markers detectable through personal consumer devices and online social media. The findings are expected to identify gaps in the field of digital phenotyping and to orient recommendations toward further research into the potential of smartphones, consumer wearables and online social media. The results will be disseminated through publication and applicable conferences. As the methodology followed includes evaluation of publicly available material, there is no need for ethical approval.

Twitter Alex Bottle @DrAlexBottle

Contributors PS and YR were involved in writing protocol, preliminary literature review; JC and AB were involved in editing and review of protocol. All authors have made substantive intellectual contributions to the development of this protocol and read and approved the manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval Formal ethical approval is not required, as primary data will not be collected in this study.

Provenance and peer review Not commissioned; externally peer reviewed.

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 iDs
Pier Spinazze http://orcid.org/0000-0002-1416-0174
Josip Car http://orcid.org/0000-0001-8969-371X

REFERENCES
3 IDC. Always connected how Smartphones and social keep us engaged 2013.