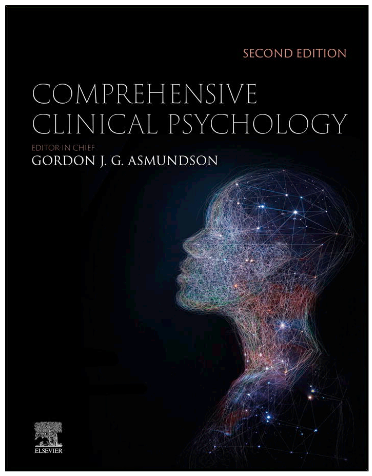

From Heinz, M.V., Thomas, N.X., Nguyen, N.D., Griffin, T.Z., Jacobson, N.C., 2022. 4.17. Technological Advances in Clinical Assessment. In: Asmundson, G.J.G. (Ed.), Comprehensive Clinical Psychology, 2nd edition, vol. 4. Elsevier, pp. 301–320.
https://dx.doi.org/10.1016/B978-0-12-818697-8.00171-0.
ISBN: 9780128186978

4.17 Technological Advances in Clinical Assessment

Michael V Heinz, Natalie X Thomas, Nhi D Nguyen, Tess Z Griffin, and Nicholas C Jacobson, Dartmouth College, Lebanon, NH, United States

Comprehensive Clinical Psychology, Second Edition, 2022, 301–320

4.17.1 Introduction

4.17.1.1 Current Assessments

At the time of this writing, psychological assessments and diagnoses are both largely guided by self-report, whether in the form of patient self-report questionnaires, such as the Patient Health Questionnaire (PHQ), or in the form of structured clinical interviews, like the structured clinical interview for the DSM-5 (SCID-5) (First, 2015). Current assessments have enabled many advancements in the field of psychology and psychiatry by standardizing diagnoses for treatment, prognosis and research. The Diagnostic and Statistical Manual, now in its fifth edition, is used by mental health professional to diagnose reliably mental disorders in order to guide treatment recommendations, identify prevalence rates for mental health services, identifying patient groups for clinical and basic research, and documenting public health information on mental disorders, such as morbidity and mortality rates (American Psychiatric Association, 2013). There is continual revision to the DSM to improve upon its clinical utility. For instance, recent updates have taken efforts to incorporate more of the neurobiological basis of mental disorders such as dissociative disorders which can help better inform assessments and subsequent treatment (Brand et al., 2012).

Many traditional self-report inventories and interviews have high reliability and validity. For example, studies have found the Generalized Anxiety Disorder Questionnaire-IV(GAD-Q-IV) to have good test-retest reliability and convergent and discriminant validity (Robinson et al., 2010). The GAD-Q-IV is valuable as both a screening tool and an indicator of the disorder's severity. Many screening and diagnostic questionnaires also exist for depression, and studies over the last several decades have attempted to compare their validity. One such study compared the validity of three questionnaires, namely Hospital Anxiety and Depression Scale (HADS), the WHO (five) Well Being Index (WBI-5), and the Patient Health Questionnaire (PHQ) (Löwe et al., 2004). The study results revealed high internal consistency and good validity for all three questionnaires, and they could be recommended for use in clinical practice to identify major depressive disorder as well as "any depressive disorder" (Löwe et al., 2004). Self-report scales are used in diagnosis of other mental health conditions, including PTSD and schizophrenia. The Trauma Symptom Inventory-2 (TSI-2) (Briere, 2011) and Posttraumatic Stress Disorder Checklist (PCL-5) (Blevins et al., 2015, p. 5) are well validated and widely used self-report measures for PTSD; the Revised Behavior and Symptom Identification Scale (BASIS-R) has shown promise used as a self-report screening and assessment in schizophrenia (Niv et al., 2007).

4.17.1.2 Current Diagnostic Challenges Inherent in the Illnesses

Given the successes of the aforementioned methods of assessment, it is worth noting that the assessment of psychiatric disorders and psychological problems still pose great challenges for both clinicians and researchers due to the nature of these disorders. Most of the disorders, as described by the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), are highly heterogeneous in clinical presentation, including their symptomatology, clinical features, time course, response to treatment, and long-term prognosis. Major depressive disorder (MDD), for example, is highly heterogeneous, with subtypes varying in severity, prognosis, and response to treatment (Howard et al., 2017; Iacono et al., 2020; Jakubovski and Bloch, 2014; Pu et al., 2018). Adding to this complexity, the underlying pathophysiology of most mental disorders remains poorly understood.

Tools such as the GAD-Q-IV, the PHQ-9, the TSI-2, the PCL-5, and the SCID-5 have contributed to the fields of psychology and psychiatry by addressing some of these challenges, creating a common language for researchers and clinicians to use. Despite the many strengths of these established methods of assessment, however, there is room for improvement. There is necessarily a subjective component to psychological assessment, as criteria for nearly all psychiatric disorders include information regarding inner experience (i.e., information which must be reported and cannot be directly observed); such assessments may be improved by supplementary objective data, collected naturalistically and in real time from devices like smartphones and wearables. An understanding of the current limitations of traditional assessments is crucial in motivating the need for the methods we propose in this chapter.

4.17.1.3 Some Limitations of Traditional Psychological Assessment

4.17.1.3.1 Subjectivity to Retrospective Biases

Emotional recall bias, inaccuracies in affective memories, is well-established in the literature (Ben-Zeev et al., 2009; Hedges et al., 1985; Lay et al., 2017; Ready et al., 2007; Robinson and Clore, 2002; Sato and Kawahara, 2011). Though diagnostic interviews and some mental health questionnaires capture current emotional state, many heavily rely on retrospective self-report, requiring the patient to access and recall emotional memory. The Patient Health Questionnaire screening for depression (PHQ-9), for instance, relies on the patient's recall over the last 2 weeks (Kroenke et al., 2001). Peak and recency biases, (Lay et al., 2017), describe the way emotional memory tends to be more weighted by either (1) recent events with emotional valence or (2) interval events with unusually high emotional valence. Emotional recall bias may also be heightened by negative mood states (Sato and Kawahara, 2011) and become even more inaccurate in pathological states such as depression (Ben-Zeev et al., 2009).

4.17.1.3.2 Subjectivity in Measurement

Questionnaire-based psychological assessments are subjective. While this is necessarily the case (psychological distress cannot be objectively quantified, and arguably the best way to know what someone feels is to ask), and suggests that assessment is dependent on the fidelity and the accuracy of memory, recall, and insight. Complicating this further, individuals with psychiatric illnesses, like depression and social anxiety, may have distorted cognitions (Johnson et al., 1992), potentially skewing their reports of interval information.

4.17.1.3.3 Obtrusiveness and Burden

Both self-assessment questionnaires and clinical interviews (structured interviews based on standardized diagnostic criteria) are obtrusive, as they require time and recall effort on the part of the patient. Clinical interviews come with the added burden of requiring employment of a mental health professional; in the US this is both a limited resource and a potential monetary cost for the patient.

4.17.1.3.4 Inability to Detect Important Time-Dependent Patterns

Most predominant methods of psychological assessment rely on a patient's memory and summary of interval events and associated timelines. Questionnaires and clinical assessments are not administered continuously, but rather at discrete time points, often separated by long and irregular time intervals. The nature of traditional assessments, therefore, limits their ability to model time-dependent patterns or to detect causal relationships.

4.17.1.3.5 Inability to Account for Simultaneous Contextual or Ecological Contributing Factors

Like their challenges in capturing time-dependent information, traditional assessment measures are also limited by an inability to objectively account for contextual and environmental factors associated with psychiatric symptoms. While severe or high intensity environmental changes (such as the death of a loved one or being fired from a job) may be remembered and therefore detected on traditional clinical assessments, more subtle, time-dependent symptom patterns associated with environmental influences may be missed. Take for example, moderate increases in anxiety associated with talking on the phone, or going to the supermarket, which are not consciously remembered, but which cause clinically-detectable increases in psychiatric symptoms.

4.17.1.3.6 Inability to Integrate With Simultaneous Physiological or Behavioral Phenomena

Traditional psychiatric and psychological assessment methods are not particularly well-suited to take into account naturalistic physiological changes associated with pathologic states. They instead rely on a patient's capacity to detect, remember, and later recall changes, possibly subtle, in her physiological state, and often over prolonged periods of time.

4.17.1.4 Potential for Improvements on the Horizon

With recent advances in computing and increasing accessibility to large amounts of data ("big data"), there is the potential for use of these data to supplement traditional psychiatric and psychological assessments. Large datasets are particularly useful in building computational or mathematical models that have the potential to enhance traditional clinical assessment, and these models can be used to make inferences and predictions regarding prognosis, and treatment of mental disorders. Although the focus of this chapter is on the applications of machine learning and not the technical aspects of its methods, some technical information is needed to assist the reader in fully understanding its application.

The first part of this chapter introduces concepts important toward understanding the more technical aspects, including foundational information with regard to machine learning models central to the topics discussed in this chapter. Fundamentally, a model is a mathematical function that takes a set of inputs, often called features, changes them in some way, and produces a set of outputs. In order for the model to be useful and to accurately reflect relationships of interest in data, it is *trained* on existing data with known inputs and outputs. Model training involves an iterative process of fine changes to the model (changes to what are referred to as model parameters), such that it approximates patterns and relationships between inputs and outputs in our data sample. The training process can be thought of as a calibration or "fitting" of the model to a particular data sample. Typically, the model is then validated by making sure it maintains the ability to model similar relationships within unseen data from the sample. This process of validation helps to ensure that the model has not *overfit* to the training data. Overfitting in this context refers to a model that has been too closely fit its training sample, with consequent diminished capacity to generalize to new data (i.e., data which has not been used in training). Such a model "misses the forest for the trees," likened to a student who misses "big picture" concepts by focusing too heavily on contextualized details. It is from these ideas that *machine learning* follows. Machine learning models refer specifically to models which can "learn" implicitly from training data, without explicit programming by humans. Such models ideally generalize well to new, related data and may elucidate potentially complex relationships in data, with the goal of better understanding reality.

For example, consider a dataset comprising demographic data (i.e., age, gender, etc.) and a binary suicide outcome (i.e., suicide, no suicide) from high risk, suicidal patients seen in a regional emergency department over three years. Let us now imagine we train a model with such data, iteratively allowing it to learn implicitly from the training examples. We would, in effect, be creating a simplified representation of reality, specifically the complex factors that impact suicide risk. After training such a model, we might use it to either predict suicide risk for new patients who come in to the emergency department (i.e., using this model to support clinical decision making), or we might investigate the inner workings of the model (this is possible with complex algorithms, a discussion of which is beyond the scope of this chapter) to determine which factors are most important in informing suicide risk. In this way, such models can be used to generate, empirical, potentially novel suicide risk factors.

In the following chapter, we focus primarily on studies which predict outcomes constituting binary variables (i.e., suicide vs no suicide), categorical variables (i.e., ultra-high risk, high risk, intermediate risk, low risk), continuous variables (i.e., percent likelihood that a particular patient ends her life).

In a clinical context, model inputs might include patient age, gender, and presence of substance use (as determined by a drug screen, for instance), and a model output might be a patient's likelihood of suicide in the coming year. Such a model would be trained on past data where both inputs and outputs were already known, and the trained model could then be used to make inferences about whether a new patient is likely to end her life, given a certain set of clinical features. *Features* in this context refer to those elements which would be used as inputs to the model for prediction, and feature sets can vary greatly in size and type.

At this point the reader may be asking the question, "now that we have these models, what's next and what do these mean for clinical assessment at present."

Use of computational models to inform healthcare is currently an emerging area of research. Models trained on mental health data have the potential to be used in a variety of ways: such models (1) may be integrated into electronic medical records (EMRs) to provide clinical decision support, (2) may uncover novel digital biomarkers[1] with the potential to aid in clinical assessment, (3) and additionally may serve as "proof of concept" for larger scale technologies that can be developed to inform clinical assessment. Machine learning models in healthcare are trained with anonymized aggregate data to protect patient privacy. In the remainder of this chapter, we will discuss novel methods for clinical assessment made possible by large datasets including those derived from electronic medical records, mobile passive sensing data streams, social media platforms, and search engine data.

4.17.2 Electronic Medical Record Data in Clinical Assessment

4.17.2.1 Topic Overview

Electronic medical records (EMRs) are used in a variety of clinical settings, from outpatient medical offices to inpatient units, and over the last 20 years, there has been a major growth in use of EMRs in clinical settings. Electronic medical records allow for the collection of large amounts of patient data; in the US, 96% of all hospitals and 86% of office-based practices had adopted use of an EMR (The Office of the National Coordinator for Health Information Technology, 2017). In addition, nearly 30% of the total stored data in the world is estimated to be generated by the healthcare industry (Huesch and Mosher, 2017), including both **structured** and **unstructured** data (**Table 1**) from both mental health and non-mental health clinical encounters.

Electronic medical record data include patient demographics, comorbid health conditions, allergies, labs, medications, clinical narratives, physical and mental status examinations, imaging, and treatment outcomes, among many other features. Structured data, data such as age and smoking status that are easily stored in discrete, structured formats within an EMR, require very little preprocessing prior to use in a computational model, and lend well to tabular display format. Unstructured data, data including clinical narratives and imaging (**Table 1**), however, remains relatively untapped by researchers despite it constituting over 80% of EMR data (Kong, 2019). These data require significant preprocessing prior to use, which includes extracting discrete, meaningful features. Consider our example above training a model to predict suicide. We might consider broadening our dataset to include unstructured clinical notes to provide our model with additional information to improve its performance. As a simple example, one such feature we might extract from such clinical notes would be the number of times the word "death" appears in the text.

Table 1 Electronic medical record data types

Examples of structured data	*Examples of unstructured data*
Race/ethnicity	Narrative in clinical note
Age (ordinal)	Assessment and plan in clinical note
Smoking status (binary)	Descriptive results of psychological testing
Weight (continuous)	Raw imaging data from MRI

[1]The term *digital biomarker* may be unfamiliar to some readers. It refers to particular user generated data patterns (perhaps collected in an electronic medical record or social media platform) that have significant association with clinical outcomes of interest. An example of a biomarker for depression might be reduced motion and decreased ambient light, detected by a mobile phone's sensors.

This discrete numerical data could then be used as an input to our model, perhaps improving prediction. In addition to input data, output data are also often available in an EMR, and can include binary features related to presence or absence of disease of interest, time to onset of illness, and prognosis after illness onset, among other data.

Structured data refer to data like age, gender, and race, which are typically stored in discrete, *structured* format in an electronic health medical record. Such data conform well to a tabular format and do not require much preprocessing. This is in contrast to unstructured data, which include free-text clinical narrative notes and imaging saved in the electronic medical record. Such data typically require much preprocessing to extract useful features that can be used to train models.

4.17.2.2 Use of EMR Data in Clinical Assessment of Suicide

Clinical assessment of suicide risk is notoriously difficult, even for clinicians with expertize in psychological assessment (Capodanno and Targum, 1983). McDowell et al. (2011) found that one month prior to suicide, nearly 50% of all individuals who commit suicide visit their primary care physician, and 20% visit a psychiatrist (McDowell et al., 2011). One large obstacle in suicide prediction is the rarity of this event, a challenge mitigated by availability of large datasets such as EMRs. Use of EMR data to train machine learning models has shown promise in accurately predicting suicide risk (Ben-Ari and Hammond, 2015; Poulin et al., 2014). This research has made use of both unstructured text from clinical notes (McCoy et al., 2019; Poulin et al., 2014) as well as structured EMR features, including number of mental health visits and patient age (Ben-Ari and Hammond, 2015). Poulin et al. using natural language processing (NLP)[2] analysis of clinical notes written by mental health providers about provider-patient interactions trained a model to 65% or greater predictive accuracy. This suggests the possibility of using such a model for screening of asymptomatic patients or suicide risk stratification. Use of computational models trained on EMR data in suicide risk research shows potential in use of routinely collected clinical biomarkers, as well as clinical notes to identify patients who may be at elevated risk, and can help clinicians provide early, targeted intervention. Fig. 1 shows a hypothetical model, which could be used to infer individual suicide risk. This model includes both (1) structured input clinical features, such as age and gender, as well as (2) unstructured input features, such as those extracted from clinical narratives.

4.17.2.3 Use of EMR Data in Clinical Assessment of Depression

Similar machine learning models have shown promise in assessment of depressive disorders. Depression poses challenges to both researchers and clinicians because of its heterogeneity in presentation, prognosis, and treatment response, and to date, several studies have attempted to address these challenges. A 2017 study (Pham et al., 2017) developed a machine learning model capable

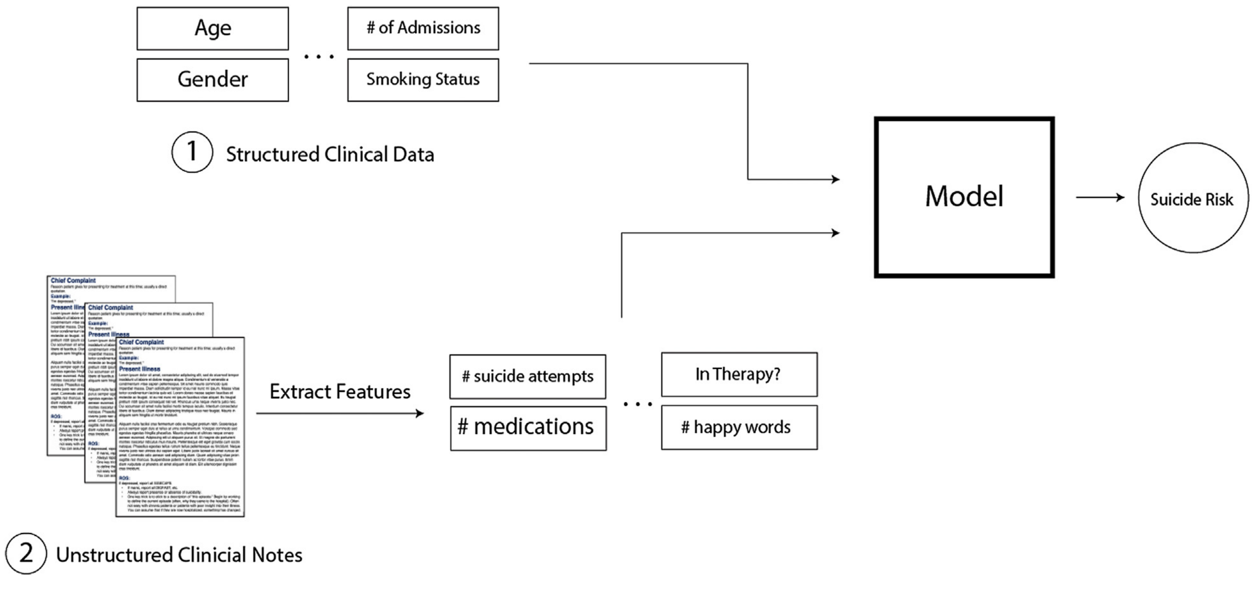

Fig. 1 Possible future applications of such a model, where routinely collected data is used to make inferences about suicide risk, has the potential to both identify individuals who need further screening and to stratify patients according to risk.

[2]NLP refers to a variety of techniques used to extract structured features from human generated, unstructured text. NLP covers a broad range of processes, including simple ones, like extracting meta features (like average word count over a body of text) to intermediate complexity (like emotionally valent word counts in a body of text) to very complex statistical models, which infer linguistic context.

of encoding information across time, using clinical notes from hospital admissions across time. The model in this study made use of Long Short Term Memory (LSTM) networks, machine learning models capable of accounting for event dependencies across time. It is worth thinking about LSTMs in the context of major depressive disorder, given the longitudinal nature of this illness, as well as its long and short term dependencies, which may help inform clinical assessment, treatment, and prognosis. Similarly, other studies have used EMR data to help assess depressive disorders, including use of structured and unstructured EMR data to predict depression severity and response to treatment (Huang et al., 2014), use of unstructured clinical narratives to predict presence of depression (Tran and Kavuluru, 2017), and use of structured EMR data to quantify the impact of medical conditions on depression (Ryu et al., 2016). Perlis et al. (2012) used clinical concepts extracted from EMRs to classify patients as either depressed or not depressed, and found that use of unstructured clinical notes proved more accurate than use of billing codes toward prediction of depression. Utilizing no information explicitly related to mental health, a recent study made use of biomedical and demographic features collected at routine college health screenings to make accurate predictions about depressive illness (as well as anxiety) (Nemesure et al., 2021). The authors identified variables including blood pressure, heart rate, and weight to inform these predictive models of depression and anxiety.

4.17.2.4 Use of EMR Data in Clinical Assessment of Schizophrenia and PTSD

NLP techniques have also proved successful in identifying negative symptoms of schizophrenia, including apathy, decreased motivation, and social withdrawal, in the electronic medical record (Gorrell et al., 2013; Patel et al., 2014, 2015). Negative symptoms have been of particular interest, as they can serve as predictors of worse clinical outcomes for schizophrenic patients. These studies are important not just because they use automated methods to identify patients with negative symptoms, but because they take advantage of the large dataset offered by the EMR to identify associated factors which may not be obvious otherwise. For example, associations have been found between negative symptoms of schizophrenia and each of the following: younger age, male gender, and increased likelihood of hospital admission (Patel et al., 2014, 2015). Such findings would not be possible without a very large, high fidelity, feature-rich data set, like that provided by an EMR.

Schultebraucks et al. (2020) found that, along with common PTSD predictors, additional, seemingly unrelated data in the EMR were valuable in prediction. The authors found that clinical features including neutrophil and lymphocyte count, blood glucose, creatinine, age, and blood pressure all had value in predicting non-remitting PTSD in these patients. Interestingly, and much like Nemesure et al. these features had no evident direct relationship to traditional diagnostic assessments, and would not have been readily evident otherwise.

4.17.2.5 Use of EMR Data Toward Identification of Novel Biomarkers

An advantage to machine learning methods using large, feature-rich datasets is their ability to potentially identify novel biomarkers associated with psychiatric disorders. Take, for example, the aforementioned EMR studies conducted by Nemesure et al. and Patel et al. These studies highlighted novel associations between features and illness without reliance on a priori assumptions or previously well-established clinical risk factors. Recall Nemesure et al. identified variables including blood pressure, heart rate, and weight to exhibit a significant association with anxiety and depression; this result is important as it draws attention to such factors not well established, even in official diagnostic criteria, that may be important markers of pathophysiology and subsequently the diagnosis of psychological illnesses. Similarly, Patel et al. found features like male gender and increased likelihood of hospital admission to be related to negative symptoms of schizophrenia, associations difficult to establish without a large, high-fidelity dataset.

Identification of novel biomarkers has direct impact in both clinical and translational research domains. In a clinical setting, discrete biomarkers may be used to inform clinical judgment for more accurate psychological assessment. EMR studies may also serve as "proof of concept" for future development of integrated clinical decision support systems. Such automated systems would use machine learning methods with EMR data inputs to provide timely recommendations to clinicians based on a patients' unique EMR data. For example, a patient with EMR data including impulsive behavior, past suicide attempts, and guns in the home, might trigger a notification to a clinician that the patient is at high risk for suicide, and might benefit from additional support measures. In a research domain, identification of biomarkers may prompt novel research questions, or spur new areas of research with later translation to clinical care.

4.17.2.6 Use of EMR Data in Predicting Psychiatric Inpatient Outcomes

EMR data have also proven useful at predicting outcomes of interest for psychiatric inpatients, like the need for seclusion to maintain safety (Hazewinkel et al., 2019). The authors of this study used NLP techniques to find language markers for patients who were likely to require seclusion. The NLP in this study entailed mining EMR text for words or concepts which had differing frequencies between medical records of patients who did require seclusion compared to those patients who did not. Hazelwinkel et al. found that in addition to words and concepts that were more obviously indicative of seclusion, including "seclusion" and "threatening," so too were words not so readily apparent in their ability to predict a seclusion treatment outcome of interest, including "cigarettes," "office," "radio," "door," and "ground," among others.

McCoy et al. have also made important contributions to the study of use of EMRs in psychological assessment and diagnosis. Their work examined the use of NLP techniques to extract meaningful patient information across RDoC symptom domains. RDoC, short for the Research Domain Criteria Initiative, was developed by the National Institute of Mental Health (NIMH) to serve as a research framework for investigating mental disorders across multiple symptoms domains (Development of the RDoC Frame-Work., n.d.). As noted by the NIMH, the RDoC framework is not intended to be a diagnostic framework, but rather a framework for research. The RDoC comprises six transdiagnostic domains (negative valence, positive valence, cognitive systems, social process systems, arousal and regulatory systems, and sensorimotor systems), and each can be investigated across *units of analysis*. These include genes, molecules, cells, circuits, physiology, behavior, and subjective report. For example, a study on sustained threat (a construct within the negative valence domain) might investigate the behavioral unit of analysis: avoidance patterns. Multiple studies by McCoy et al. have helped establish the validity of RDoC symptom domains in clinical contexts using machine learning (McCoy et al., 2019; McCoy and Perlis, 2015). Their 2015 study found significant associations between length of inpatient psychiatric hospital stay and both cognitive and arousal domains; their 2018 study found significant positive correlations between length of stay and cognitive and arousal domains as well as negative correlations between positive and negative affect domains and length of stay.

As we conclude the first section of this chapter (focused on the application of machine learning to EMR data), it is worth general consideration of the potential benefits high computing power and machine learning bring to the psychology when applied to large passively collected datasets. Machine learning combined with these data make possible more personalized screening and assessment, as they supplement traditional self-report methods by integrating specific, highly dimensional and feature rich data. Such an approach has the potential to personalize clinical screening and assessment by accounting for more person specific data previously untapped (e.g., EMR vitals, labs, and narratives). As we shall see in the following section, mobile and wearable devices allow for integration of naturalistic physiologic and behavioral supplements to assessment.

4.17.3 Mobile Device Passive Sensing Data in Clinical Assessment

4.17.3.1 Overview

Fundamentally, passive sensing data are data collected without the active involvement of the user. Many mobile devices currently on the market are replete with sensors with the ability to gather feature-rich datasets unique to an individual. The most recent iteration of Apple's iPhone, the iPhone 11, is equipped with many sensors, including a barometer (for measuring pressure), a gyroscope (for measuring rotational motion), an accelerometer (for measuring motion), a proximity sensor (which can detect the closeness of an object to the device), an ambient light sensor, as well as a microphone and multiple cameras. The iPhone is also equipped with a GPS system, Bluetooth, and WiFi capabilities. Current market competitors, including the Samsung Galaxy and the Google Pixel, have similar sensors and capabilities. Also growing in popularity are wrist-worn devices that serve as extensions to smartphones, which can also come equipped with a number of sensors including accelerometers and heart rate monitors.

Smartphone sensors and what they measure

Sensor type	*What does sensor measure?*
Barometer	Pressure
Gyroscope	Rotational acceleration
Accelerometer	Linear acceleration
Proximity sensor	Phone's distance from an object (often used to darken)

Additional data sources on smartphones

Incoming texts or calls
Outgoing texts or calls
Global positioning system (GPS)
Connection or disconnection to WiFi network
Time

When combined with the ubiquity of smartphone use, information gathered from these sensors allows for the collection of large amounts of real-time data, even from a single user. In addition, many phone companies also make these sensors available to third party application developers, enabling developers to build applications that collect and use these data. Often these apps route these data to servers where the data are preprocessed and used to train machine learning models. This movement of data from a mobile device to a server is often referred to as a **passive sensing stream**.

Passive sensing streams can result in the collection of huge amounts of data, and a growing number of applications available for download onto mobile devices now take advantage of various sensor data. The result is an unequaled ability to gain large amounts of information on a single individual, including information that can impact that person's mental health assessment.

4.17.3.2 Fictional Case Example

Take, for example, a fictional character named Sue. Sue has undiagnosed social anxiety disorder and mild major depressive disorder; she lives alone and works for a local bank. Sue owns both a smartphone and a smart wrist-worn device, and has recently installed an application on her phone that she's hoping can help her with the growing stress she feels at work. This app runs in the background on her phone, and without any effort on her part collects de-identified information about her in real-time. Through use of Sue's passive sensing data stream, this app can make inferences about her psychological condition, and subsequently provide Sue with customized mental health recommendations. Accelerometer data, potentially combined with light sensor data and the clock, would allow this application to make inferences about Sue's sleep patterns over time. Microphones detecting de-identified human speech in the background could provide information on the frequency and duration of Sue's in-person interactions with co-workers. Call logs or phone proximity sensors would allow for inferences about how much time Sue spends talking on the phone. Accelerometer and gyroscopic sensors could provide information on Sue's physical activity level throughout the day. In addition, a paired wrist worn device might give information about Sue's pulse, and paired with ambient light and noise sensors might inform her physiologic response to environments with high levels of noise.

In Sue's case example, it's clear that much useful information can be collected from mobile device sensors, and combined (for example, combining heart rate sensor and ambient noise sensor) can lead to even more powerful and complex behavioral features. Collected over time, these large amounts of data describing Sue's day-to-day behavior might even be used to begin to make inferences about her thoughts and emotions. This longitudinal data and the patterns therein, are described as an individual's **digital phenotype,** a term coined by Torous et al. in 2015 as "moment-by-moment quantification of the individual-level human phenotype in-situ using data from smartphones and other personal digital devices." Growing literature in the mental health field makes use of passive sensing data streams.

4.17.3.3 Mobile Passive Sensing Data in Behavioral Assessment of Mood Disorders

4.17.3.3.1 Bipolar Disorder

Current literature suggests the possibility of identification of bipolar disorder mood episodes (i.e., depressed or manic), as well as transition periods between such states through use of passive data streams. In research conducted by Grunerbl et al. correlations were found between motion, location, and phone call data with both bipolar state and transition periods between states (Grüenerbl et al., 2012, 2014, 2015). Other studies investigating passive sensing in bipolar patients have found predictive ability in features suggestive of social interaction (i.e., number of incoming calls, outgoing calls, and text messages) as well as movement, as indicated by connectivity to cell phone towers (Faurholt-Jepsen et al., 2014, 2015, 2016a,b). Additional research suggests an impressive ability (AUC of 0.89 and 0.78 for manic and depressed states, respectively) of voice analysis in predicting manic and depressed mood states (Faurholt-Jepsen et al., 2016a,b). These findings are in agreement with known behavioral manifestations of bipolar illness, including increased social activity and increased locomotion.

4.17.3.3.2 Depressive Disorders

Research has also indicated great promise in the ability of passive data streams to be used toward predicting depressive disorders. The Dartmouth StudentLife Study, a 10-week study developed to better understand mental health disorders in the college student population (Wang et al., 2014), passively collected smartphone data to test the ability of such data to predict mental health outcomes, including depressive disorders. The authors found a robust correlation (absolute r values in the range of 0.3–0.4) between passively sensed behaviors and depression (as gauged by PHQ-9 scores), with sleep, conversation frequency and duration, and participant location of particular importance in mental health predictive ability. It is worth noting that self-report measures are often used as the ground truth or target metric for model prediction. While these self report instruments have limitations (discussed at the beginning of this chapter), their sensitivity and specificity for categorical psychiatric illness (such as the PHQ for Major Depressive Disorder) is moderate to high. This allows for studies such as StudentLife to establish convergent validity between categorical diagnoses and passively collected data, an important step in furthering digital mental health research. Additionally, such methods allow for passively collected data to complement existing diagnostic frameworks. The StudentLife study also used objective outcome measures of functioning, like cumulative and term grade point averages.

The StudentLife Study demonstrates the power in combining individual sensor data types to create high-level behavioral features. For example, the authors index higher level features, including "activity," "conversation," "sleep," and "locomotion" using various combinations of accelerometer, microphone, light, GPS, and Bluetooth sensors. These higher level behavioral features were then used to make mental health inferences.

To date similar research has used personal, mobile passive sensing data streams to make inferences about depression. Much of this research has relied heavily on either movement features, location features, or a combination of the two (Canzian and Musolesi, 2015; Chow et al., 2017; Farhan et al., 2016; Jacobson et al., 2019a; Lu et al., 2018; Saeb et al., 2016; Ware et al., 2018; Yue et al., 2017). Given the potential for missing GPS data, authors of some studies used create solutions to optimize longitudinal location monitoring, reducing the potential for missing data. Yue et al. for example supplement GPS data with WiFi data to reduce potential for missing location data and find improved predictive results. Ware et al. rely exclusively on WiFi data to infer user location and use this to predict depression. Other studies have made use of passive light data (Jacobson et al., 2019b) and audio data (Rabbi et al., 2011), each combined with movement data to make inferences about depression, as well as communication data (Boukhechba et al., 2018) and interactivity data (Mehrotra et al., 2016), including incoming calls and texts and phone usage data to make such inferences. Intuitively, these studies generally reflect current knowledge about behavioral manifestations of depression. Depressed people are more likely to have reduced activity, reduced locomotion, and reduced social engagement.

While most of the aforementioned studies address depression as a single time point outcome, promise has been shown in the ability of passive sensing data to predict moment-to-moment changes in depressive symptoms over time (Jacobson and Chung, 2020). The research shows promise in addressing the question of how depression symptoms change in hour to hour time intervals. The results of this study suggest the possibility of "just in time" treatment interventions, customized interventions provided to users during nearly real-time detected symptom changes.

4.17.3.4 Mobile Passive Sensing Data in Behavioral Assessment of Psychotic Disorders

Passive sensing data, like EMR data, have also been used to make inferences regarding symptom severity, trajectory, and relapse in patients with schizophrenia. Despite possible challenges regarding tolerability of passive sensors in this population, Ben-Zeev et al. have shown that passive sensing is feasible and generally well-tolerated among patients with schizophrenia (Ben-Zeev et al., 2015). A 2016 study assessing mental health changes in individuals with schizophrenia found that passively-collected smartphone data could predict general mental health state in patients. Higher levels of physical activity were associated with more positive affect, and higher levels of virtual communication (calls and texts) were associated with increased negative affect measures (Wang et al., 2016).

The onset and relapse of psychotic symptoms (sometimes referred to as decompensation) in patients with schizophrenia have traditionally been very difficult to predict with self-report measures (Halpin and Carr, 2000). To date, multiple studies utilizing passive sensing data address this difficulty, showing positive results in predicting relapse (Barnett et al., 2018; Buck et al., 2019; Wang et al., 2020), most often using a 30-day time window prior to relapse. Depp et al. found that GPS data could be successfully used to predict negative symptom domains in schizophrenia, and Wang et al. found that symptom trajectories in patients recently discharged from the hospital could be predicted by passive sensing data.

Albeit to a lesser extent than bipolar disorder and depressive disorders, other domains of psychological pathology have also been studied using passive sensing data. To date, multiple passive sensing studies have examined the association between anxiety disorders and passive data, including movement (gauged by accelerometer) and social contact (gauged by incoming calls and texts) data (Boukhechba et al., 2018; Jacobson et al., 2020b). Boukhechba et al. examined GPS data and, notably, found a significant correlation between time spent out of town and increased social anxiety, (among other increases in negative affect). Similar to the StudentLife study, Fukazawa et al. combined data involving acceleration, app usage, phone angle to identify, for example, whether a participant was sitting at a desk looking at their phone or walking while using a phone application (Fukazawa et al., 2019). These behaviors were combined and used as features to make higher-level construct inferences about anxiety changes over time. Even studies using single passive sensing data streams have found significant predictive use of passive sensing data toward making mental health predictions. A study by Jacobson et al. found predictive success in predicting worry using only actigraphy (movement) data in HIV patients (Jacobson and O'Cleirigh, 2019).

4.17.4 Social Media Data in Clinical Assessment

4.17.4.1 Topic Overview

Social media platforms have become increasingly popular within the last several decades, including Facebook, Twitter, LinkedIn, Instagram, Snapchat, Reddit, and many more. These networks are commonly used for social interaction, relaxation, communication, expression of opinion, and information sharing (Whiting and Williams, 2013), and due to their wide array of features, social media platforms can provide a relatively intimate look at an individual's personality traits and potentially psychopathology through that individual's interactions with these platforms. These online behaviors and actions can be collected and processed in unobtrusive, accurate, and efficient ways (Hao et al., 2014; Zhu et al., 2016), and prior research has investigated the utility of using social media platforms in behavioral research, including studying personality (Kosinski et al., 2013; Schwartz et al., 2013) and subjective well being (Hao et al., 2014). Such studies have been successful in predicting psychologically related outcomes by examining variables including social media posts, profiles, and status updates.

4.17.4.2 Popular Social Media Platforms

4.17.4.2.1 Twitter

A social media platform in which individuals share opinions and thoughts in short posts of up to 140 characters called "tweets". This platform is of particular use in its ability to spread information quickly and to a wide audience. Twitter currently has 330 million monthly active users, 145 million daily active users, and 500 million tweets are sent per day (10 Twitter Statistics Every Marketer Should Know in 2020 [Infographic], 2019).

4.17.4.2.2 Facebook

Facebook is a social networking site in which an individual can create a profile account with personal photographs, life events, updates, demographic information, and text posts. This platform is especially useful in its ability to connect friends and family and as a way of sharing information, and currently has 2.7 billion monthly active users (Facebook, n.d.).

4.17.4.2.3 Reddit

Reddit is a social media site that emphasizes discussion and community. Reddit users can consume content by reading articles posted to the forum, can create content (of up to 40,000 characters), can interact with content by rating and commenting, and have the ability to interact with one another through communities. and According to (Shen and Rudzicz, 2017) there are 234 million Reddit users that post around 75.15 million posts and 725.85 million comments a month. Additionally, Reddit has sub-pages called subreddits that center around unique topics, and are utilized by individuals to create niche communities, seek advice, and to share information and personal experiences.

4.17.4.2.4 Instagram

Instagram is a social networking platform that focuses on visual media. Each post consists of some visual aid, such as a picture, and an optional text caption. Users create profiles, post content, and create a following of other users, based on their posts. Instagram is a particularly popular platform for individuals to promote certain ideas and products among their followers.

4.17.4.3 Social Media Data as a Window to Personality Characteristics

Prior research using social media data has focused on its ability to help us learn more about personality traits (Finch and Graziano, 2001). Researchers have found that social exchanges and web behavior on social media platforms can provide information on certain personality characteristics such as agreeableness, extraversion, and neuroticism as well as an individual's interpersonal characteristics including social support and social exchange, thus potentially informing an individual's mental health status (Finch and Graziano, 2001). Use of social media in mental health assessment has subsequently become an area of growing research.

4.17.4.4 Social Media Data Types

Like electronic medical records discussed earlier in this chapter, social media includes both structured and unstructured data. Recall that structured data lends itself well to tabular format and requires little preprocessing prior to modeling. In a social media context, structured data could include the number of posts in a given time period and user demographics, like age and gender. To inform a model about substance use, for instance, we might consider the number of posts by participants in a given time period and the age and gender of the user. To gain additional information, we might use unstructured social media components, which could include content of user posts. These data types, while rich in information, require substantial preprocessing prior to use in a model.

NLP techniques are useful in extracting features from unstructured content which may have the ability to inform clinical assessment (**Fig. 1**). For example, an individual post can be broken down to look at specific language used, including word choice or diction, pronoun choice, length of post, number of and number of swear words. While the topic or content of the post itself may reflect something as mundane as the weather, the language used can inform an individual's mental health status. For example, a study by Reece & Danforth examined the validity of using Instagram pictures and visual content to predict depression (Reece and Danforth, 2017). They found that users who tended to use darker colors in their pictures were later diagnosed with depression. In this instance, while the actual image itself, considered an unstructured data point, did not necessarily indicate a user's state of depression, an individual component of their image, color, did. Unstructured data from individual social media posts can also be analyzed to find overall trends in specific mental health conditions (Guntuku et al., 2017).

4.17.4.5 Useful Features of Social Media Platforms

Despite a multitude of social media platforms, they all have a few defining features that make them highly useful in clinical assessment. All of these platforms collect personal and longitudinal data from a wide range of users that are readily accessible. By doing so, social media data is particularly useful toward understanding how mental health symptoms change over time or the **trajectory** of a specific individual's mental disorder. As discussed earlier, mental disorders such as depression are extremely heterogeneous in nature, and as a result, predicting or even understanding an individual's prognosis can be difficult. However, by looking at an individual's data over time, a much better idea of their particular mental health symptoms and progression can be formed (De Choudhury, 2013). Much of the existing research on social media and mental health has capitalized on this aspect of social media data in order to predict individual trajectories of mental health conditions, and in turn, understand overall patterns of mental health. Similar to many of the other data types discussed in this chapter, social media data is also useful in creating models to better understand and predict onset and other features of a specific mental disorder. For example, a study by Choudhury et al. used crowdsourcing techniques to gather individual Twitter data from clinically depressed users. These data were used as input to train machine learning models used to predict depression onset for particular individuals (Choudhury et al., 2013). Study researchers also found overall trends in social media use associated with depression.

> Mental health conditions largely display implicit changes in language and behavior, such as a switch in the types of topics, a shift in word usage or a shift in frequency of posts.
>
> Coppersmith et al. (2014).

4.17.4.6 Use of Social Media in Predicting Mental Health

In the past, mental health status has been largely determined by medical records, gold standard clinical assessments, and answers to behavioral surveys. While these methods remain critical to clinical assessment, they can be obtrusive, burdensome and subject to biases. Use of social media in conjunction with gold standard clinical assessment can lessen some of those burdens. One of the most promising features of passive data such as social media data is its ability to detect implicit changes in behaviors, a characteristic of certain mental health conditions, and can include subtle shifts in language, topics, and frequency of posts as reflected in social media data (Coppersmith et al., 2014). Additionally, social media data is captured in a naturalistic setting and in real-time, giving it a temporal advantage that can detect small patterns and changes in behavior that infrequent assessments might not capture. These data are akin to a "digital diary" of an individual's life, capturing posts, birthdays, status updates and more. Real-time data also helps alleviate subjectivity to retrospective biases. And as content has already been generated, these data do not require obtrusive or burdensome assessments in order to collect meaningful information. The nature of these data also reflect contextual or ecological factors that may impact an individual's mental health status including interactions with friends, community posts, and public forums, and provide a closer look at an individual's interests and life events.

4.17.4.6.1 Predicting Depression

Researchers have used social media data to train machine learning models to distinguish healthy individuals from non-healthy individuals, and in doing so have discovered trends in mental health disorders including depression, anxiety disorders, mood disorders, and schizophrenia. With just two months of observational data as input, these models have been shown to accurately predict individual disorder trajectories as well as overarching patterns and trends in mental health (Choudhury et al., 2013; Eichstaedt et al., 2018; Hu et al., 2015). In addition, social media data has the potential to help predict early symptoms of mental disorders, as well as prevent their onset.

One of the most notable studies using Twitter data examined the capacity of such data to characterize the onset of depression. In this study (Choudhury et al., 2013), used crowdsourcing to collect a years' worth of Twitter posts from users diagnosed as depressed as indicated by the Center for Epidemiologic Studies Depression Scale (CESD), prior to depression onset. Researchers measured several different factors including engagement, use of depression language, emotion, linguistic style, and egocentric social graph, defined as a social interaction measure. Based on these measures, a machine learning model was trained to classify twitter user's as either depressed or non-depressed. This study found differences in timing, volume, social connection, and language for depressed and non-depressed individuals, and to date, more current research exhibits similar trends in timing, language, and social connection as classifiers for a variety of mental health disorders (Guntuku et al., 2017; Park et al., 2013; Whiting and Williams, 2013). This study is an example of how social media data can be used in conjunction with traditional assessments to predict mental health outcomes.

4.17.4.6.2 Time Stamps as a Mental Health Assessment Variable

Understanding timing, progression, and trajectory of a mental health disorder is crucial in clinical assessment. Social media data provide time stamps that can offer added insight into specific timing and onset of symptoms related to mental health disorders. For example, users with mental health disorder symptoms tend to be more active during night hours and less active during the day (Choudhury et al., 2013). This social media behavior may be indicative of insomnia or lack of sleep, a possible symptom of depression. Social media data also allows researchers to understand when mental health symptoms are more likely to occur. Schwartz et al. found that predictive models for depression using Facebook posts were more accurate in the winter, indicating higher degrees of depression in the wintertime (Schwartz et al., 2014). Both timing and time stamps of an individual's social media behavior can indicate behavioral symptoms related to mental health disorders, and clinicians can look for these patterns in their patients to determine risk or likelihood of developing certain mental health conditions.

4.17.4.6.3 Social Connection and Volume of Posting as Mental Health Assessment Variables

Social withdrawal and lack of interest are key diagnostic indicators of depression, which can be shown through social media data, thus providing insight into social connectedness and interests of users. An individual's social interactions with others can reveal a good deal of information about their mental health status, and multiple studies have looked at users' online social interactions and behaviors including number of followers, instant messaging habits, re-tweets, and comments as indicators of mental health (Park et al., 2013; Williams and Galliher, 2006). For example, lack of social media activity and interaction with others could signify mental health symptoms, and might be inferred by volume of posts, frequency of posting, interactions with friends, and number of friends (Birnbaum et al., 2017; Choudhury et al., 2013; De Choudhury et al., 2016; Park et al., 2013; Ricard et al., 2018). A study by Park et al. (2013) using Facebook data, found that a user's number of friends and location tags were negatively correlated with their depressive scores as indicated by the CESD (Park et al., 2013). Ricard et al. (2018) explored the possibility of using community-generated content instead of user-generated content to identify depression among social media users. Community-generated content included data generated by a user's friends or followers, such as comments, likes and number of followers (Ricard et al., 2018). Their study found that community-generated content was largely complementary to user-generated content in predicting user depression, and demonstrated the potential for a user's social network and community to also provide insight into that individual's mental health status. Williams & Galliher demonstrated similar findings in the college student population (Williams and Galliher, 2006), and found that, among college students, lowered levels of social competence and connectedness were related to higher levels of depression.

In addition to depression, withdrawal from social connection is also prevalent in other mental health disorders. In a study analyzing 671 Twitter posts, Birnbaum et al. found that users with schizophrenia demonstrated increased use of first person pronouns, de-emphasized friendships, and seemed to talk more about biological processes than the non-psychosis twitter users (Birnbaum et al., 2017). In this research, biological processes refer to language usage including words such as "body" or "health", indicating an increased focus and awareness of oneself. These results reflect many of the abovementioned studies indicating that social media data can depict mental health states based on language usage.

4.17.4.6.4 Language Used in Social Media as a Mental Health Assessment Variable

Passive sensing data such as social media data can detect implicit changes in behavior. Usually this involves extracting meaningful information from unstructured data, and data in social media is found in user post text. Multiple studies have used machine learning models to process language used in users' social media accounts (Choudhury et al., 2013; Coppersmith et al., 2015; De Choudhury et al., 2016; Eichstaedt et al., 2018; Schwartz et al., 2014). A hallmark of depressive disorders is rumination on sustained, negative self-thoughts (Choudhury et al., 2013). In 2013 Choudhury et al. found that individuals displaying a greater use of first person pronouns, swear words, and depressed language were later more likely to be diagnosed with depression (Choudhury et al., 2013). This pattern has also been found in other mental health disorders including suicidal ideation, schizophrenia, and anxiety (De Choudhury et al., 2016; Eichstaedt et al., 2018; Schwartz et al., 2014; Shen and Rudzicz, 2017).

4.17.4.6.5 Language Usage in Anxiety and PTSD

Increased interpersonal pronoun use has also been found to be associated with other mental health disorders. A study by Reece et al. found that Twitter data can be used to predict depression and PTSD, as the two are often comorbid (Reece et al., 2017) Another study found statistically significant language differences in Twitter data between individuals self-diagnosed with PTSD and randomly-selected individuals. Most striking in this work was the finding that tweets from users with self-diagnosed PTSD contained more third-person pronouns and language regarding anxiety (Coppersmith et al., 2014). Additional research has looked into detecting anxiety using Reddit posts. In 2017, Shen & Rudzicz analyzed 22,808 Reddit posts (9971 anxiety-related posts and 12,837 general posts) collected over the course of three months (Shen and Rudzicz, 2017). A model was created to predict anxiety onset using social media data, and they found results similar to those looking at depression onset. Anxious posts contained more first-person singular pronouns and focused on topics related to anxiety, drugs, and school. This work suggests the potential for social media data scraping to help identify topics related with certain mental health disorders, and understanding trends in language usage among users with mental health disorders may serve useful in detecting mental health disorder cases.

4.17.4.6.6 Predicting Suicidality

Certain patterns in language use, frequency and timing of posts, and social connection can also indicate suicidality in a user. Suicide is one of the top ten leading causes of death in the US, resulting in 1.4% of total adult deaths (De Choudhury et al., 2016). Suicidal ideation is difficult to diagnose and to detect as these individuals are less likely to outwardly show symptoms (De Choudhury et al., 2016). Similar to depression, suicidality presents itself differently depending on the individual. It is therefore important to find ways to detect individuals displaying symptoms of suicidality, particularly among those who may not explicitly or outwardly indicate suicidality. A study on suicide analyzing Reddit data, created predictive models using three categories: linguistic structure (nouns, verbs, and adverbs), interpersonal awareness (pronoun use), and interaction (volume and length of posts, volume and length of comments on the posts). This study found that users at risk for suicidality used increased levels of first person pronouns, showed decreased levels of social engagement, and had a decreased overall post readability (De Choudhury et al., 2016). These overall patterns in behavior are similar to those found in depression research.

4.17.4.6.7 Predicting Bipolar Disorder

Social media platforms, such as Facebook, provide users an outlet with which to express their current emotion and vent frustrations (Eichstaedt et al., 2018), suggesting this platform's potential for tracking changes in a user's emotion over time. Several psychopathologies, including depressive disorders, center on the inability to regulate emotions and are marked by dysfunctional abilities to emotionally regulate negative thoughts pertaining to the self (Choudhury et al., 2013). In 2018, Seabrook et al. investigated whether machine learning models could predict not only the presence of mental health symptoms, but changes in these symptoms as well. As a major component of depression is marked by a high degree of emotion variability (Seabrook et al., 2018), found that greater negative emotion word instability to be indicative of increased depression severity.

Emotion phasic mood disturbances and mood instability are also prominent characteristics in bipolar disorder. As social media data has the ability to reflect changes in mood over time, these data may demonstrate great utility in predicting onset of bipolar disorder. One study analyzed Twitter data to predict possible signs of the initial prodrome stage of bipolar disorder. This study used a novel time-specific subconscious crowdsourcing approach in order to identify the time period of bipolar disorder onset in Twitter users (Huang et al., 2017). This research relied on both, the assumption that social media users subconsciously reveal information, such as mental health status, in their social media activity, as well as on crowdsourcing to gather information from a large group of people (Huang et al., 2017). Huang et al.'s model looked at trends such as insomnia and over-talking to predict onset of bipolar disorder with 91% accuracy (Huang et al., 2017). A study by Saha et al. analyzed Twitter posts to predict and infer mood instability (defined as rapid and unpredictable shifts in mood from baseline to depression, irability, or anxiety). Their study showed potential in use of social media data in predicting bipolar disorder among Twitter users (Saha et al., 2017). Saha et al. created a model that could predict mood instability with 96% accuracy, and that could predict self-identifying users with bipolar disorder and borderline personality disorder, as indicated by higher levels of mood instability (Saha et al., 2017). Another study found that Reddit posts can predict bipolar disorder based on the number of emotion-expressive words (Sekulić et al., 2019). The ability of social media data to track longitudinal changes in emotional language makes it ideal for use in predictive models of bipolar disorder.

4.17.4.7 Potential Limitations

Multiple studies recommend use of social media predictive models as screening tools in conjunction with more traditional clinical assessments (Choudhury et al., 2013; De Choudhury et al., 2016; Ernala et al., 2019). Labeling an individual as at risk for suicidal ideation can have lasting effects on their self-perception, well-being, and overall health, and it's important to have multiple methods of confirming this diagnosis. Furthermore, data from social media research may not be generalizable to the global population. For example, social media data taken from specific users in a Reddit subgroup may not be generalizable to a greater population. Additionally, these prediction models work better with more data, and in general, users with less social media data receive a less-personalized output from prediction models when social media data alone is used in prediction. It is important to note that the accuracy of machine learning models relies on a multitude of social media data types, and on data of sufficient quantity to generate predictions based on past behaviors, and that when used in conjunction with other data types described in this chapter, will ultimately result in a more accurate model for mental health prediction.

4.17.5 Search Engine Data in Clinical Assessment

4.17.5.1 Topic Overview

As individuals increasingly rely on the internet as a source of health information, search data has become more widely utilized as a data source in clinical assessment. In this context, search data refers to information collected on internet search engine usage, including search engines Google, Bing, and Yahoo. In 2009, it was estimated that between "37% to 52% of Americans (sought out) health-related information on the Internet each year, generally using search engines to find advice on conditions, symptoms, and treatments" (Brownstein et al., 2009). Google is currently the most popular search engine worldwide, and, as a result, existing research has often focused on using Google search data to predict health risk and potential outbreaks (Lee, 2020). In particular, Google Trends search volume is most often used as a way to track the daily, weekly, and monthly popularity of search terms of interest (Lee, 2020). Monitoring time-sensitive trends in search data can provide an early indicator of health risk within a population and help inform subsequent clinical assessment.

4.17.5.2 Advantages Over Traditional Assessment

Search data's biggest advantage over traditional methods of clinical assessment and surveillance is the ability to collect this type of data frequently. In clinical assessment, this is especially important as it allows for identification of key time-dependent patterns of certain pathologies and does not rely on a patient's single retrospective recall. Additionally, search data collection isn't obtrusive or burdensome to the individual, as the data are collected without any interference into the patient's everyday life. Furthermore, search data is also objectively collected and without susceptibility to bias from either the patient or the clinical professional.

4.17.5.3 Historic Roots: Search Engine Data in Epidemic Surveillance

While applying search data in clinical settings is still relatively new, early studies in the field demonstrated success with using search data to track epidemics and diseases. This work has since expanded to include research on mental health issues and search data.

With seasonal influenza epidemics a major public health concern, and traditional methods of surveillance reporting lags of 1–2 weeks, Ginsberg et al. conducted a study aimed at utilizing influenza-related search data to achieve faster community detection (Ginsberg et al., 2009). In order to monitor indirect signals of influenza activity, they sought to develop a "simple model that estimates the probability that a random physician visit in a particular region is related to an influenza-like illness

(ILI)" (Ginsberg et al., 2009). They used a single explanatory variable, the probability that a random search query submitted from the same region is ILI-related, and search queries included in the model were identified using an automated method of selecting ILI-related search queries data. Ginsberg et al. found that the "relative frequency of certain queries was highly correlated with the percentage of influenza-like illness (ILI) physician visits" (Ginsberg et al., 2009). With this model, they were able to estimate the current level of weekly influenza activity in each region of the United States, with a reporting lag of about one day. This proved a significant improvement compared to traditional assessments. This decrease in time lag advantage of search data proves especially crucial when looking at certain psychopathologies. Although application has been much more limited in tracking mental-health related issues, the successful application of search data in detecting influenza epidemics can serve as an indicator of the potential for search data in clinical assessment.

At the time of this writing, a recently published 2020 study (Jacobson et al., 2020a) looked at the relationship between stay-at-home orders for the COVID-19 epidemic, and the impact of such measures on mental health. The study analyzed over 10 million Google search queries across the United States. Study results found that mental health related search queries increased prior to the issue of stay-at-home orders. However, after the orders were enacted, there was a leveling off in search queries, especially those related suicidal ideation and anxiety, after the orders were enacted (Jacobson et al., 2020a). This is of interest considering limited social contact is expected to have a negative effect on mental health. While further research is necessary to examine more long-term effects of governmental measures during a pandemic, this study successfully utilized search data to investigate the impact of epidemics on mental health, and demonstrated the potential of search data in mental health research.

4.17.5.4 Search Engine Data in Mental Health

Most studies using search data toward mental health surveillance and assessment involve identifying key terms related to a pathology (i.e., suicide, depression) and examining the relationship between search volume of these key terms and real-life incidences of these pathologies. The aim with this work is to use these relationships toward early prediction and detection of health trends, and hopefully enable appropriate intervention at community and individual levels. As an emerging area of research, there is limited literature available, often focusing on certain pathologies, and success has varied geographically.

4.17.5.4.1 Suicide

Search data have been used most heavily in studies on suicide. Suicide risk assessment poses a challenge for clinical professionals for a variety of reasons including its being a rare event and its heterogeneity in nature. Additionally, data collection on suicide is hampered by time lags as "public health statistics are often released too late to affect reversible societal factors affecting suicide" (McCarthy, 2010). Yang et al. argue that "even with standardized measurement of suicide risks, recognition of suicidal individuals is still difficult and preventive intervention is often too late." Considering these limitations, use of search data in suicide research has garnered attention due to its ability to reduce this time lag. In addition, search data use is specifically useful in this population, as there is a growing use of the internet for "persons with suicidal intentions [who] may not seek medical/psychiatric attention, but rather seek information regarding the means of suicide" (Yang et al., 2011). This shift in behavior has placed even more importance on analyzing search data, as it can provide particular insight into this population and help with providing them appropriate clinical assessments.

There have been several key studies that have capitalized on these advantages and used search data in suicide assessment. Conducted between 2004 and 2009, one study (McCarthy, 2010) looked at Google search engine activity for suicide-related terms. In particular, search volumes of the terms "suicide", "teen suicide", "depression", "divorce" and "unemployment" were tracked using Google Trends. The study found that Google search volumes correlated with CDC statistics for both suicide and self-injury exhibiting patterns that differed by age (McCarthy, 2010). A second study, took a similar approach but instead focused on suicide death data in Taipei City, Taiwan (Yang et al., 2011). In contrast to McCarthy, this study analyzed 37 suicide-related terms representing major known risks of suicide, including psychiatric, medical, familial, socioeconomic factors, and pro-suicide terms (Yang et al., 2011). This study found that "searches for most medical, familial, and socioeconomic terms preceded suicide deaths and most searches for psychiatric related terms coincided with suicide data." Both of these studies show the potential of search data application in suicide surveillance and intervention as demonstrated by correlations between suicide-related search terms and real-life incidences of suicide deaths. Researchers in this area are also adopting new methods to improve predictions in order to most accurately track pathologies and better inform early public health intervention. A recent study by Lee (2020) used machine learning models to improve predictive ability, in averaging multiple time series per search term.

A common trend in studies applying search data to suicide research is a limited clinical application of their findings, specifically the correlations found between suicide-related search terms and real-life instances of suicide deaths. However, in a study conducted from 2011 to 2015 (Chai et al., 2019) researchers utilized relationships between specific search terms and suicide deaths in order to develop an early warning system of suicide in Hong Kong. While this detection system has encountered various challenges, including a small population size and a relatively reduced suicide incidence rate, this detection system has been crucial in informing timely public health interventions, with highest accuracy demonstrated in the <25 year age group (Chai et al., 2019). More research is needed to fully utilize the predictive power of suicide-related search data in clinical detection and assessment.

4.17.5.4.2 Depression

While search data has mainly been used in suicide research, it is gradually expanding to other mental health disorder research areas, including depression. In the past, traditional methods of data collection for depression have looked at trends on a broad scale, typically tracking annual and seasonal variations (Tana et al., 2018). As a result, traditional data has remained incomplete toward "the study of diurnal mood variations, or daily changes in individuals' overall mood state in relation to depression-like symptoms" (Tana et al., 2018). With time serving as an important factor in clinical assessment of depression, use of search data in assessment can offer major advantages over more traditional methods. Tana et al. have argued that search data are especially useful as "some of the temporal variations and clock-like rhythms that govern several different health-related behaviors can be traced in near real-time with the help of search engine data".

A key study in this field, Tana et al. (2018) used search data to monitor depressive symptoms by "analyzing diurnal variations for interest in depression on the Web to discover hourly patterns of depression interest and help seeking." In this particular study, the following six depression-related search queries (in Finnish) were tracked: "Masennus" (Eng. depression), "Masennus oireet" (Eng. depression symptoms), "Masennustesti" (Eng. depression test), "Masennus testi" (Eng. depression test), "Depression," and "Depression test," the English and Sweedish terms for depression and depression test (Tana et al., 2018). Tana et al. found that "help seeking for depression has clear diurnal patterns, with significant rise in depression-related query volumes toward the evening and night" (Tana et al., 2018). This constant flow of data collection using search data that isn't subject to biases or gaps enables a near real-time tracking of patients' mood and depression-like symptoms that traditional data fails to capture. These findings prove important as understanding these trends can help inform health professionals in providing temporally-appropriate interventions to patients.

4.17.5.5 Future Directions for Search Engine Data in Mental Health

At present, data collected from sources such as Google Trends remain at a societal and aggregate level. These empirically derived data provide valuable information about the biomarkers and risk factors associated with various disorders. Given the strengths of using search engine data in mental health surveillance and assessment, there is still room for improvement in this area of research. Future research could potentially focus on collecting and analyzing more individualized data to obtain more personalized assessments, while still ensuring protection of personal user information. Given a specific user's search data, the goal would be to utilize that information to provide a more accurate indicator of health risk for that specific individual, and inform more personalized interventions. Efforts would need to be made to balance a more individualized collection of data with maintaining a person's privacy.

Additionally, search data have only been narrowly applied in clinical settings, with most research in this area focused on suicide. With its great predictive potential, future directions may include expanding search data use to surveillance of other mental health disorders. Future research could also include finding added ways to utilize the predictive capabilities of search data, in order to obtain more direct clinical application and impact, with use of search data to develop an early detection system in Hong Kong providing insight into what that might look like.

4.17.6 Conclusions and Limitations

Millions of people are affected by mental health issues every year with at least 1 in 5 adults experiencing some sort of mental health disorder every year (Mental Health By the Numbers and NAMI: National Alliance on Mental Illness, n.d.). Mental health disorders can be heavily burdensome for individuals, families, and communities (Patel et al., 2016). The economic burden of mental illnesses was an estimated 2.5 trillion globally in 2010. This number is expected to double within the next two decades (Whiteford et al., 2013). There is evidence of prolonged time delays between symptom onset, diagnosis, and treatment (Altamura et al., 2010). The cause of these delays is multifactorial and may include: (1) the mental health provider shortage (worldwide), (2) the lack of insight often present in conditions with egosyntonic features, like OCD (Alonso et al., 2008), and (3) provider misclassification (Vermani et al., 2011) perhaps linked to known underreporting of ofttimes stigmatized symptoms (Bharadwaj et al., 2015). Factor (1) may be improved by technologies which make possible scalable assessments, and factors (2) and (3) may be addressed by technologies which complement present diagnostic measures by incorporating objective behavioral and physiologic data.

Diagnosis of mental health disorders poses many challenges, which traditional diagnostic assessments are not fully equipped to handle. Mental illnesses represent a highly heterogeneous group of disorders, complex and variable in their cognitive and behavioral phenotypes. Traditional assessments rely on broad categorical diagnostic frameworks, which are unable to fully model the complexity of these illnesses.

Recent advances in large data collection, storage, and processing (made possible by increased computational power) have allowed for the development of more sophisticated, computerized models of such mental illnesses. Machine learning models, trained with large datasets, allow for a finer, more nuanced understanding of these illnesses and the complex interplay between neural circuitry, behavior, and the environment. Various channels, including electronic medical records, mobile passive sensing streams, social media, and search engines all provide large datasets, with mental health-related features, which have utility in developing, less obtrusive, scalable systems for mental health assessment. This chapter aims to provide the reader with a high level overview of emerging areas of research, relying on such channels.

4.17.6.1 Electronic Medical Record Data and Associated Studies

Current EMRs contain vast amounts of stored biomedical data, both structured and unstructured. Studies to date have found significant and clinically meaningful associations between both structured and unstructured components. Natural language processing has been a useful tool in extracting structured variables from unstructured data for use in machine learning models. Many of these models have found significant associations between EMR data and mental health outcomes.

4.17.6.2 Passive Sensing Data Streams and Associated Studies

Passive sensing entails data collected without the active involvement or participation of the patient. Mobile devices provide a nearly ubiquitous platform, equipped with many passive sensors, for collection of large amounts of data which have been found by recent studies to be reflective of behavioral and cognitive phenotypes.

4.17.6.3 Social Media Data and Associated Studies

Social media data can pick up on a variety of behaviors associated with a person's mental health status. With a vast majority of people documenting their behaviors on social media platforms, they provide large amounts of personalized longitudinal data collected in real time, thus giving valuable insight into a person's mental health status through their interactions with others, posts, language use, and other internet behaviors.

4.17.6.4 Internet Search Data and Associated Studies

Search data allows for timely detection and tracking of broad mental health trends without intrusion into the patient's life. With a growing population using the internet for health-related information, the amount of data available is vast and robust. Studies have found that mental health trends derived from search data follow real-life instances of these pathologies. Thus, this predictive power offered by training machine learning models with search data provides potential for forming systems of early detection and informing the appropriate public health interventions.

The data collected from all of the above channels have several common features, which make them well suited for our task of improving mental health assessment.

(1) They provide very large amounts of individualized data, useful for training robust machine learning models capable of prediction and inference.
(2) They provide objective measures of an individual's past behaviors; hence, they are not reliant on often faulty retrospective memory.
(3) They do not require active participation on the part of the user or deviation from their usual workflow; hence, they are relatively unobtrusive, compared to traditional assessments.
(4) Data collection is longitudinal in nature, making possible elucidation of time or environmental dependencies, often lost in traditional assessment.

Machine learning models combined with large, passively collected datasets combined with current self-report methods have potential for improving individual and population level mental health. On an individual level, we can consider time dependent models which use naturalistic, longitudinal, moment to moment mobile data to make predictions about real-time mental health outcomes. We might imagine such a model having utility for providing "just in time" interventions (i.e., interventions which are not only tailored to an individual, but also customized to an individual's in-the-moment mental health needs) based on real time assessments. As an example, recall the use of passively collected data to predict moment to moment changes in depression severity (Jacobson and Chung, 2020). We might also imagine a model which can match a patient to optimal treatment, based on individual-level factors, which may include biodemographic data or behavioral and physiologic data.

On a population level, creating these personalized learning models can act as a scalable, low cost, method for clinical assessment. As this data is unobtrusive and widely available, it can easily be scrapped and used to inform machine learning models. Scalable and widely deployed predictive models could act as screening measures for mental illnesses. In doing so, they could (a) provide early intervention for any persons exhibiting symptoms, known to be associated with improved trajectories (Allen et al., 2007), and (b) potentially prevent the relapse of serious mental illnesses, like schizophrenia. This could significantly lower the number of people who go untreated for their mental health disorders, while reducing the costs of clinical assessment. Overall, looking at technology as an accessible, and widespread clinical assessment tool for mental health disorders can have long last outcomes.

Lastly, we will discuss limitations and pitfalls to these emerging areas of clinical assessment.

(1) Self-report measures are often used as the ground truth or target metric for model prediction. While these self-report instruments have limitations (many of which are discussed at the beginning of this chapter), their sensitivity and specificity for categorical psychiatric illness (such as the PHQ for Major Depressive Disorder) is moderate to high. Rather than considering technological assessments a replacement for subjective, self-report metrics, we recognize the complementarity of objective,

passively collected data to subjective report. When conceptualized in this manner, we may imagine beginning to bridge the gap between internal, subjective experience and associated external, behavioral manifestations.

(2) Complex machine learning models, even those with good predictive ability, may **lack interpretability**, making it difficult to ascertain which features have the greatest importance in driving model predictions. Nemesure et al. addressed this with the use of novel software capable of quantifying relative feature importances.
(3) Machine learning models are subject to **overfitting**, meaning that they overfit to the data they are trained on, but do not perform well on unseen, new data. This can lead to erroneous inferences about the population of interest.
(4) Identifying features with high predictive ability in large datasets can lead to false conclusions of causal relationships between such features and outcomes of interest.
(5) Clear paths do not exist for immediate clinical implementation of mental health models. EMR-integrated decision support systems provide one possibility, but at this time are not widely used.
(6) Machine learning models may reinforce or propagate group stereotypes through identification of such features as racial, ethnic, or socioeconomic group status as predictive of mental illness.
(7) There are potential ethical issues in collecting large amounts of data regarding a subject matter as sensitive as mental health. When conducting this type of research, it is important to consider a few different factors: user expectations, user privacy, and researcher responsibility (Conway and O'Connor, 2016).
(8) Machine learning model predictions of high risk/high liability outcomes may pose ethical and practical problems in real-life clinical contexts. For example, suppose a model trained on suicide EMR and passive sensing data predicts a high likelihood of suicide in a patient sitting in front of us in the emergency department who is denying suicide. What do we do next?

Nevertheless, these technological advances represent an impending potential paradigm shift in their ability to gather continuous behavioral and physiological assessments that are directly tied to mental health.

References

10 Twitter Statistics Every Marketer Should Know in 2020 [Infographic], (2019). https://www.oberlo.com/blog/twitter-statistics.

Allen, N.B., Hetrick, S.E., Simmons, J.G., Hickie, I.B., 2007. Early intervention for depressive disorders in young people: the opportunity and the (lack of) evidence. Med. J. Aust. 187 (S7). https://doi.org/10.5694/j.1326-5377.2007.tb01329.x.

Alonso, P., Menchón, J.M., Segalàs, C., Jaurrieta, N., Jiménez-Murcia, S., Cardoner, N., Labad, J., Real, E., Pertusa, A., Vallejo, J., 2008. Clinical implications of insight assessment in obsessive-compulsive disorder. Compr. Psychiatry 49 (3), 305–312. https://doi.org/10.1016/j.comppsych.2007.09.005.

Altamura, A.C., Buoli, M., Albano, A., Dell'Osso, B., 2010. Age at onset and latency to treatment (duration of untreated illness) in patients with mood and anxiety disorders: a naturalistic study. Int. Clin. Psychopharmacol. 25 (3), 172–179. https://doi.org/10.1097/YIC.0b013e3283384c74.

American Psychiatric Association, 2013. Diagnostic and Statistical Manual of Mental Disorders, fifth ed.

Barnett, I., Torous, J., Staples, P., Sandoval, L., Keshavan, M., Onnela, J.P., 2018. Relapse prediction in schizophrenia through digital phenotyping: a pilot study. Neuropsychopharmacology 43 (8), 1660–1666. https://doi.org/10.1038/s41386-018-0030-z.

Ben-Ari, A., Hammond, K., 2015. Text Mining the EMR for Modeling and Predicting Suicidal Behavior Among US Veterans of the 1991 Persian Gulf War, 2015 48th Hawaii International Conference on System Sciences, pp. 3168–3175. https://doi.org/10.1109/HICSS.2015.382.

Ben-Zeev, D., Young, M.A., Madsen, J.W., 2009. Retrospective recall of affect in clinically depressed individuals and controls. Cogn. Emot. 23 (5), 1021–1040. https://doi.org/10.1080/02699930802607937.

Ben-Zeev, D., Wang, R., Abdullah, S., Brian, R., Scherer, E.A., Mistler, L.A., Hauser, M., Kane, J.M., Campbell, A., Choudhury, T., 2015. Mobile behavioral sensing for outpatients and inpatients with schizophrenia. Psychiatr. Serv. 67 (5), 558–561. https://doi.org/10.1176/appi.ps.201500130.

Bharadwaj, P., Pai, M.M., Suziedelyte, A., 2015. Mental Health Stigma. Working Paper No. 21240; Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w21240.

Birnbaum, M.L., Ernala, S.K., Rizvi, A.F., De Choudhury, M., Kane, J.M., 2017. A collaborative approach to identifying social media markers of schizophrenia by employing machine learning and clinical appraisals. J. Med. Internet Res. 19 (8), e289. https://doi.org/10.2196/jmir.7956.

Blevins, C.A., Weathers, F.W., Davis, M.T., Witte, T.K., Domino, J.L., 2015. The posttraumatic stress disorder checklist for DSM-5 (PCL-5): development and initial psychometric evaluation. J. Trauma Stress 28 (6), 489–498. https://doi.org/10.1002/jts.22059.

Boukhechba, M., Daros, A.R., Fua, K., Chow, P.I., Teachman, B.A., Barnes, L.E., 2018. DemonicSalmon: monitoring mental health and social interactions of college students using smartphones. Smart Health 9–10, 192–203. https://doi.org/10.1016/j.smhl.2018.07.005.

Brand, B.L., Lanius, R., Vermetten, E., Loewenstein, R.J., Spiegel, D., 2012. Where are we going? An update on assessment, treatment, and neurobiological research in dissociative disorders as we move toward the DSM-5. J. Trauma Dissociation 13 (1), 9–31. https://doi.org/10.1080/15299732.2011.620687.

Briere, J., 2011. Trauma Sympton Inventory-2 (TSI-2) manual. Psychol. Assess. Resour. https://www.parinc.com/Products/Pkey/464.

Brownstein, J.S., Freifeld, C.C., Madoff, L.C., 2009. Digital disease detection—harnessing the web for public health surveillance. N. Engl. J. Med. 360 (21), 2153–2157. https://doi.org/10.1056/NEJMp0900702.

Buck, B., Scherer, E., Brian, R., Wang, R., Wang, W., Campbell, A., Choudhury, T., Hauser, M., Kane, J.M., Ben-Zeev, D., 2019. Relationships between smartphone social behavior and relapse in schizophrenia: a preliminary report. Schizophr. Res. 208, 167–172. https://doi.org/10.1016/j.schres.2019.03.014.

Canzian, L., Musolesi, M., 2015. Trajectories of Depression: Unobtrusive Monitoring of Depressive States by Means of Smartphone Mobility Traces Analysis. Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, pp. 1293–1304. https://doi.org/10.1145/2750858.2805845.

Capodanno, A.E., Targum, S.D., 1983. Assessment of suicide risk: some limitations in the prediction of infrequent events. J. Psychosoc. Nurs. Ment. Health Serv. 21 (5), 11–14.

Chai, Y., Luo, H., Zhang, Q., Cheng, Q., Lui, C.S.M., Yip, P.S.F., 2019. Developing an early warning system of suicide using Google Trends and media reporting. J. Affect. Disord. 255, 41–49. https://doi.org/10.1016/j.jad.2019.05.030.

Choudhury, M.D., Gamon, M., Counts, S., Horvitz, E., 2013. Predicting Depression Via Social Media. Seventh International AAAI Conference on Weblogs and Social Media. https://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/view/6124.

Chow, P.I., Fua, K., Huang, Y., Bonelli, W., Xiong, H., Barnes, L.E., Teachman, B.A., 2017. Using mobile sensing to test clinical models of depression, social anxiety, state affect, and social Isolation among college students. J. Med. Internet Res. 19 (3). https://doi.org/10.2196/jmir.6820.

Conway, M., O'Connor, D., 2016. Social media, big data, and mental health: current advances and ethical implications. Curr. Opin. Psychol. 9, 77–82. https://doi.org/10.1016/j.copsyc.2016.01.004.

Coppersmith, G., Harman, C., Dredze, M., 2014. Measuring Post Traumatic Stress Disorder in Twitter. Eighth International AAAI Conference on Weblogs and Social Media. https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8079.

Coppersmith, G., Dredze, M., Harman, C., Hollingshead, K., 2015. From ADHD to SAD: Analyzing the Language of Mental Health on Twitter Through Self-Reported Diagnoses. Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, pp. 1–10. https://doi.org/10.3115/v1/W15-1201.

De Choudhury, M., Kiciman, E., Dredze, M., Coppersmith, G., Kumar, M., 2016. Discovering Shifts to Suicidal Ideation From Mental Health Content in Social Media. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI Conference, 2016, pp. 2098–2110. https://doi.org/10.1145/2858036.2858207.

De Choudhury, M., 2013. Role of Social Media in Tackling Challenges in Mental Health. Proceedings of the 2nd International Workshop on Socially-Aware Multimedia, pp. 49–52. https://doi.org/10.1145/2509916.2509921.

Development of the RDoC Frame-Work, (n.d.). National Institute of Mental Health. Retrieved from. https://www.nimh.nih.gov/research/research-funded-by-nimh/rdoc/index.shtml.

Eichstaedt, J.C., Smith, R.J., Merchant, R.M., Ungar, L.H., Crutchley, P., Preoţiuc-Pietro, D., Asch, D.A., Schwartz, H.A., 2018. Facebook language predicts depression in medical records. Proc. Natl. Acad. Sci. U. S. A. 115 (44), 11203–11208. https://doi.org/10.1073/pnas.1802331115.

Ernala, S.K., Birnbaum, M.L., Candan, K.A., Rizvi, A.F., Sterling, W.A., Kane, J.M., De Choudhury, M., 2019. Methodological Gaps in Predicting Mental Health States From Social Media: Triangulating Diagnostic Signals. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–16. https://doi.org/10.1145/3290605.3300364.

Facebook: Active Users Worldwide, (n.d.). Statista. Retrieved from. https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/.

Farhan, A.A., Lu, J., Bi, J., Russell, A., Wang, B., Bamis, A., 2016. Multi-view Bi-clustering to Identify Smartphone Sensing Features Indicative of Depression, 2016 IEEE First International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE), pp. 264–273. https://doi.org/10.1109/CHASE.2016.27.

Faurholt Jepsen, M., Vinberg, M., Frost, M., Debel, S., Margrethe Christensen, E., Bardram, J.E., Kessing, L.V., 2016b. Behavioral activities collected through smartphones and the association with illness activity in bipolar disorder. Int. J. Methods Psychiatr. Res. 25 (4), 309–323. https://doi.org/10.1002/mpr.1502.

Faurholt-Jepsen, M., Frost, M., Vinberg, M., Christensen, E.M., Bardram, J.E., Kessing, L.V., 2014. Smartphone data as objective measures of bipolar disorder symptoms. Psychiatr. Res. 217 (1–2), 124–127. https://doi.org/10.1016/j.psychres.2014.03.009.

Faurholt-Jepsen, M., Vinberg, M., Frost, M., Christensen, E.M., Bardram, J.E., Kessing, L.V., 2015. Smartphone data as an electronic biomarker of illness activity in bipolar disorder. Bipolar Disord. 17 (7), 715–728. https://doi.org/10.1111/bdi.12332.

Faurholt-Jepsen, M., Busk, J., Frost, M., Vinberg, M., Christensen, E.M., Winther, O., Bardram, J.E., Kessing, L.V., 2016a. Voice analysis as an objective state marker in bipolar disorder. Transl. Psychiatry 6, e856. https://doi.org/10.1038/tp.2016.123.

Finch, J.F., Graziano, W.G., 2001. Predicting depression from temperament, personality, and patterns of social relations. J. Pers. 69 (1), 27–55. https://doi.org/10.1111/1467-6494.00135.

First, M.B., 2015. Structured clinical interview for the DSM (SCID). In: The Encyclopedia of Clinical Psychology. American Cancer Society, pp. 1–6. https://doi.org/10.1002/9781118625392.wbecp351.

Fukazawa, Y., Ito, T., Okimura, T., Yamashita, Y., Maeda, T., Ota, J., 2019. Predicting anxiety state using smartphone-based passive sensing. J. Biomed. Inf. 93, 103151. https://doi.org/10.1016/j.jbi.2019.103151.

Ginsberg, J., Mohebbi, M.H., Patel, R.S., Brammer, L., Smolinski, M.S., Brilliant, L., 2009. Detecting influenza epidemics using search engine query data. Nature 457 (7232), 1012–1014. https://doi.org/10.1038/nature07634.

Gorrell, G., Roberts, A., Jackson, R., Stewart, R., 2013. Finding Negative Symptoms of Schizophrenia in Patient Records. Proceedings of the Workshop on NLP for Medicine and Biology Associated with RANLP 2013, pp. 9–17. https://www.aclweb.org/anthology/W13-5102.

Grüenerbl, A., Oleksy, P., Bahle, G., Haring, C., Weppner, J., Lukowicz, P., 2012. Towards Smart Phone Based Monitoring of Bipolar Disorder. Proceedings of the Second ACM Workshop on Mobile Systems, Applications, and Services for HealthCare, pp. 1–6. https://doi.org/10.1145/2396276.2396280.

Grüenerbl, A., Osmani, V., Bahle, G., Carrasco, J.C., Oehler, S., Mayora, O., Haring, C., Lukowicz, P., 2014. Using Smart Phone Mobility Traces for the Diagnosis of Depressive and Manic Episodes in Bipolar Patients. Proceedings of the 5th Augmented Human International Conference, pp. 1–8. https://doi.org/10.1145/2582051.2582089.

Grüenerbl, A., Muaremi, A., Osmani, V., Bahle, G., Ohler, S., Tröster, G., Mayora, O., Haring, C., Lukowicz, P., 2015. Smartphone-based recognition of states and state changes in bipolar disorder patients. IEEE J. Biomed. Health Inform. 19 (1), 140–148. https://doi.org/10.1109/JBHI.2014.2343154.

Guntuku, S.C., Yaden, D.B., Kern, M.L., Ungar, L.H., Eichstaedt, J.C., 2017. Detecting depression and mental illness on social media: an integrative review. Curr. Opin. Behav. Sci. 18, 43–49. https://doi.org/10.1016/j.cobeha.2017.07.005.

Halpin, S.A., Carr, V.J., 2000. Use of quantitative rating scales to assess outcome in schizophrenia prevention studies. Aust. N. Z. J. Psychiatry 34 (s2), S150–S160. https://doi.org/10.1046/j.1440-1614.2000.00794.x.

Hao, B., Li, L., Gao, R., Li, A., Zhu, T., 2014. Sensing subjective well-being from social media. In: Ślęzak, D., Schaefer, G., Vuong, S.T., Kim, Y.S. (Eds.), Active Media Technology. Springer International Publishing, pp. 324–335. https://doi.org/10.1007/978-3-319-09912-5_27.

Hazewinkel, M.C., de Winter, R.F.P., van Est, R.W., van Hyfte, D., Wijnschenk, D., Miedema, N., Hoencamp, E., 2019. Text analysis of electronic medical records to predict seclusion in psychiatric wards: proof of concept. Front. Psychiatry 10. https://doi.org/10.3389/fpsyt.2019.00188.

Hedges, S.M., Jandorf, L., Stone, A.A., 1985. Meaning of daily mood assessments. J. Pers. Soc. Psychol. 48 (2), 428–434. https://doi.org/10.1037/0022-3514.48.2.428.

Howard, D.M., Clarke, T.K., Adams, M.J., Hafferty, J.D., Wigmore, E.M., Zeng, Y., Hall, L.S., Gibson, J., Boutin, T.S., Hayward, C., Thomson, P.A., Porteous, D.J., Smith, B.H., Murray, A.D., Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium, Haley, C.S., Deary, I.J., Whalley, H.C., McIntosh, A.M., 2017. The stratification of major depressive disorder into genetic subgroups. BioRxiv 134601. https://doi.org/10.1101/134601.

Hu, Q., Li, A., Heng, F., Li, J., Zhu, T., 2015. Predicting depression of social media user on different observation windows. In: 2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), vol. 1, pp. 361–364. https://doi.org/10.1109/WI-IAT.2015.166.

Huang, S.H., LePendu, P., Iyer, S.V., Tai-Seale, M., Carrell, D., Shah, N.H., 2014. Toward personalizing treatment for depression: predicting diagnosis and severity. J. Am. Med. Inf. Assoc. 21 (6), 1069–1075. https://doi.org/10.1136/amiajnl-2014-002733.

Huang, Y.H., Wei, L.H., Chen, Y.S., 2017. Detection of the Prodromal Phase of Bipolar Disorder from Psychological and Phonological Aspects in Social Media. ArXiv, 1712.09183 [Cs]. http://arxiv.org/abs/1712.09183.

Huesch, M.D., Mosher, T.J., 2017. Using it or losing it? The case for data scientists inside health care. NEJM Catalyst. https://catalyst.nejm.org/doi/abs/10.1056/CAT.17.0493.

Iacono, L.L., Bussone, S., Andolina, D., Tambelli, R., Troisi, A., Carola, V., 2020. Dissecting major depression: the role of blood biomarkers and adverse childhood experiences in distinguishing clinical subgroups. J. Affect. Disord. 276, 351–360. https://doi.org/10.1016/j.jad.2020.07.034.

Jacobson, N.C., Chung, Y.J., 2020. Passive sensing of prediction of moment-to-moment depressed mood among undergraduates with clinical levels of depression sample using smartphones. Sensors 20 (12). https://doi.org/10.3390/s20123572.

Jacobson, N.C., O'Cleirigh, C., 2019. Objective digital phenotypes of worry severity, pain severity and pain chronicity in persons living with HIV. Br. J. Psychiatry 1–3. https://doi.org/10.1192/bjp.2019.168.

Jacobson, N.C., Weingarden, H., Wilhelm, S., 2019a. Digital biomarkers of mood disorders and symptom change. NPJ Dig. Med. 2. https://doi.org/10.1038/s41746-019-0078-0, 3.

Jacobson, N.C., Weingarden, H., Wilhelm, S., 2019b. Using digital phenotyping to accurately detect depression severity. J. Nerv. Ment. Dis. 207 (10), 893–896. https://doi.org/10.1097/NMD.0000000000001042.

Jacobson, N.C., Lekkas, D., Price, G., Heinz, M.V., Song, M., O'Malley, A.J., Barr, P.J., 2020a. Flattening the mental health curve: COVID-19 stay-at-home orders are associated with alterations in mental health search behavior in the United States. JMIR Ment. Health 7 (6), e19347. https://doi.org/10.2196/19347.

Jacobson, N.C., Summers, B., Wilhelm, S., 2020b. Digital biomarkers of social anxiety severity: digital phenotyping using passive smartphone sensors. J. Med. Internet Res. 22 (5). https://doi.org/10.2196/16875.

Jakubovski, E., Bloch, M.H., 2014. Prognostic subgroups for citalopram response in the STAR*D trial. J. Clin. Psychiatry 75 (7), 738–747. https://doi.org/10.4088/JCP.13m08727.

Johnson, K.A., Johnson, J.E., Petzel, T.P., 1992. Social anxiety, depression, and distorted cognitions in college students. J. Soc. Clin. Psychol. 11 (2), 181–195. https://doi.org/10.1521/jscp.1992.11.2.181.

Kong, H.-J., 2019. Managing unstructured big data in healthcare system. Healthc. Inform. Res. 25 (1), 1–2. https://doi.org/10.4258/hir.2019.25.1.1.

Kosinski, M., Stillwell, D., Graepel, T., 2013. Private traits and attributes are predictable from digital records of human behavior. Proc. Natl. Acad. Sci. U. S. A. 110 (15), 5802–5805. https://doi.org/10.1073/pnas.1218772110.

Kroenke, K., Spitzer, R.L., Williams, J.B.W., 2001. The PHQ-9. J. Gen. Intern. Med. 16 (9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x.

Lay, J.C., Gerstorf, D., Scott, S.B., Pauly, T., Hoppmann, C.A., 2017. Neuroticism and extraversion magnify discrepancies between retrospective and concurrent affect reports. J. Pers. 85 (6), 817–829. https://doi.org/10.1111/jopy.12290.

Lee, J.-Y., 2020. Search trends preceding increases in suicide: a cross-correlation study of monthly Google search volume and suicide rate using transfer function models. J. Affect. Disord. 262, 155–164. https://doi.org/10.1016/j.jad.2019.11.014.

Löwe, B., Spitzer, R.L., Gräfe, K., Kroenke, K., Quenter, A., Zipfel, S., Buchholz, C., Witte, S., Herzog, W., 2004. Comparative validity of three screening questionnaires for DSM-IV depressive disorders and physicians' diagnoses. J. Affect. Disord. 78 (2), 131–140. https://doi.org/10.1016/S0165-0327(02)00237-9.

Lu, J., Shang, C., Yue, C., Morillo, R., Ware, S., Kamath, J., Bamis, A., Russell, A., Wang, B., Bi, J., 2018. Joint Modeling of Heterogeneous Sensing Data for Depression Assessment Via Multi-task Learning. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2, pp. 1–21. https://doi.org/10.1145/3191753.

McCarthy, M.J., 2010. Internet monitoring of suicide risk in the population. J. Affect. Disord. 122 (3), 277–279. https://doi.org/10.1016/j.jad.2009.08.015.

McCoy, T.H., Perlis, R.H., 2015. A clinical perspective on the relevance of research domain criteria in electronic health records. Am. J. Psychiatry 5.

McCoy, T.H., Pellegrini, A.M., Perlis, R.H., 2019. Research domain criteria scores estimated through natural language processing are associated with risk for suicide and accidental death. Depress. Anxiety 36 (5), 392–399. https://doi.org/10.1002/da.22882.

McDowell, A.K., Lineberry, T.W., Bostwick, J.M., 2011. Practical suicide-risk management for the busy primary care physician. Mayo Clin. Proc. 86 (8), 792–800. https://doi.org/10.4065/mcp.2011.0076.

Mehrotra, A., Hendley, R., Musolesi, M., 2016. Towards Multi-modal Anticipatory Monitoring of Depressive States Through the Analysis of Human-Smartphone Interaction. Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct, pp. 1132–1138. https://doi.org/10.1145/2968219.2968299.

Mental Health by the Numbers, NAMI: National Alliance on Mental Illness, (n.d.). Retrieved from. https://www.nami.org/mhstats.

Nemesure, M.D., Heinz, M.V., Huang, R., Jacobson, N.C., 2021. Predictive modeling of depression and anxiety using electronic health records and a novel machine learning approach with artificial intelligence. Sci. Rep. 11 (1), 1980. https://doi.org/10.1038/s41598-021-81368-4.

Niv, N., Cohen, A.N., Mintz, J., Ventura, J., Young, A.S., 2007. The validity of using patient self-report to assess psychotic symptoms in schizophrenia. Schizophr. Res. 90 (1–3), 245–250. https://doi.org/10.1016/j.schres.2006.11.011.

Park, S., Lee, S.W., Kwak, J., Cha, M., Jeong, B., 2013. Activities on Facebook reveal the depressive state of users. J. Med. Internet Res. 15 (10), e217. https://doi.org/10.2196/jmir.2718.

Patel, R., Jayatilleke, N., Jackson, R., Stewart, R., McGuire, P., 2014. Investigation of negative symptoms in schizophrenia with a machine learning text-mining approach. Lancet 383, S16. https://doi.org/10.1016/S0140-6736(14)60279-8.

Patel, R., Jayatilleke, N., Broadbent, M., Chang, C.K., Foskett, N., Gorrell, G., Hayes, R.D., Jackson, R., Johnston, C., Shetty, H., Roberts, A., McGuire, P., Stewart, R., 2015. Negative symptoms in schizophrenia: a study in a large clinical sample of patients using a novel automated method. BMJ Open 5 (9), e007619. https://doi.org/10.1136/bmjopen-2015-007619.

Patel, V., Chisholm, D., Parikh, R., Charlson, F.J., Degenhardt, L., Dua, T., Ferrari, A.J., Hyman, S., Laxminarayan, R., Levin, C., Lund, C., Medina Mora, M.E., Petersen, I., Scott, J., Shidhaye, R., Vijayakumar, L., Thornicroft, G., Whiteford, H., 2016. Addressing the burden of mental, neurological, and substance use disorders: key messages from Disease Control Priorities, 3rd edition. Lancet 387 (10028), 1672–1685. https://doi.org/10.1016/S0140-6736(15)00390-6.

Perlis, R.H., Iosifescu, D.V., Castro, V.M., Murphy, S.N., Gainer, V.S., Minnier, J., Cai, T., Goryachev, S., Zeng, Q., Gallagher, P.J., Fava, M., Weilburg, J.B., Churchill, S.E., Kohane, I.S., Smoller, J.W., 2012. Using electronic medical records to enable large-scale studies in psychiatry: treatment resistant depression as a model. Psychol. Med. 42 (1), 41–50. https://doi.org/10.1017/S0033291711000997.

Pham, T., Tran, T., Phung, D., Venkatesh, S., 2017. Predicting healthcare trajectories from medical records: a deep learning approach. J. Biomed. Inf. 69, 218–229. https://doi.org/10.1016/j.jbi.2017.04.001.

Poulin, C., Shiner, B., Thompson, P., Vepstas, L., Young-Xu, Y., Goertzel, B., Watts, B., Flashman, L., McAllister, T., 2014. Predicting the risk of suicide by analyzing the text of clinical notes. PLoS One 9 (1), e85733. https://doi.org/10.1371/journal.pone.0085733.

Pu, S., Noda, T., Setoyama, S., Nakagome, K., 2018. Empirical evidence for discrete neurocognitive subgroups in patients with non-psychotic major depressive disorder: clinical implications. Psychol. Med. 48 (16), 2717–2729. https://doi.org/10.1017/S003329171800034X.

Rabbi, M., Ali, S., Choudhury, T., Berke, E., 2011. Passive and In-Situ Assessment of Mental and Physical Well-Being Using Mobile Sensors. Proceedings of the 13th International Conference on Ubiquitous Computing—UbiComp '11, p. 385. https://doi.org/10.1145/2030112.2030164.

Ready, R.E., Weinberger, M.I., Jones, K.M., 2007. How happy have you felt lately? Two diary studies of emotion recall in older and younger adults. Cogn. Emot. 21 (4), 728–757. https://doi.org/10.1080/02699930600948269.

Reece, A.G., Danforth, C.M., 2017. Instagram photos reveal predictive markers of depression. EPJ Data Sci. 6 (1), 1–12. https://doi.org/10.1140/epjds/s13688-017-0110-z.

Reece, A.G., Reagan, A.J., Lix, K.L.M., Dodds, P.S., Danforth, C.M., Langer, E.J., 2017. Forecasting the onset and course of mental illness with Twitter data. Sci. Rep. 7 (1), 13006. https://doi.org/10.1038/s41598-017-12961-9.

Ricard, B.J., Marsch, L.A., Crosier, B., Hassanpour, S., 2018. Exploring the utility of community-generated social media content for detecting depression: an analytical study on Instagram. J. Med. Internet Res. 20 (12), e11817. https://doi.org/10.2196/11817.

Robinson, M.D., Clore, G.L., 2002. Belief and feeling: evidence for an accessibility model of emotional self-report. Psychol. Bull. 128 (6), 934–960. https://doi.org/10.1037/0033-2909.128.6.934.

Robinson, C.M., Klenck, S.C., Norton, P.J., 2010. Psychometric properties of the Generalized Anxiety Disorder Questionnaire for DSM-IV among four racial groups. Cogn. Behav. Ther. 39 (4), 251–261. https://doi.org/10.1080/16506073.2010.486841.

Ryu, E., Chamberlain, A.M., Pendegraft, R.S., Petterson, T.M., Bobo, W.V., Pathak, J., 2016. Quantifying the impact of chronic conditions on a diagnosis of major depressive disorder in adults: a cohort study using linked electronic medical records. BMC Psychiatry 16. https://doi.org/10.1186/s12888-016-0821-x.

Saeb, S., Lattie, E.G., Schueller, S.M., Kording, K.P., Mohr, D.C., 2016. The relationship between mobile phone location sensor data and depressive symptom severity. PeerJ 4, e2537. https://doi.org/10.7717/peerj.2537.

Saha, K., Chan, L., De Barbaro, K., Abowd, G.D., De Choudhury, M., 2017. Inferring Mood Instability on Social Media by Leveraging Ecological Momentary Assessments. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 1, pp. 95:1–95:27. https://doi.org/10.1145/3130960.

Sato, H., Kawahara, J., 2011. Selective bias in retrospective self-reports of negative mood states. Hist. Philos. Logic 24 (4), 359–367. https://doi.org/10.1080/10615806.2010.543132.

Schultebraucks, K., Shalev, A.Y., Michopoulos, V., Grudzen, C.R., Shin, S.M., Stevens, J.S., Maples-Keller, J.L., Jovanovic, T., Bonanno, G.A., Rothbaum, B.O., Marmar, C.R., Nemeroff, C.B., Ressler, K.J., Galatzer-Levy, I.R., 2020. A validated predictive algorithm of post-traumatic stress course following emergency department admission after a traumatic stressor. Nat. Med. 26 (7), 1084–1088. https://doi.org/10.1038/s41591-020-0951-z.

Schwartz, H.A., Eichstaedt, J.C., Kern, M.L., Dziurzynski, L., Ramones, S.M., Agrawal, M., Shah, A., Kosinski, M., Stillwell, D., Seligman, M.E.P., Ungar, L.H., 2013. Personality, gender, and age in the language of social media: the open-vocabulary approach. PLoS One 8 (9), e73791. https://doi.org/10.1371/journal.pone.0073791.

Schwartz, H.A., Eichstaedt, J., Kern, M.L., Park, G., Sap, M., Stillwell, D., Kosinski, M., Ungar, L., 2014. Towards Assessing Changes in Degree of Depression Through Facebook. Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, pp. 118–125. https://doi.org/10.3115/v1/W14-3214.

Seabrook, E.M., Kern, M.L., Fulcher, B.D., Rickard, N.S., 2018. Predicting depression from language-based emotion dynamics: longitudinal analysis of Facebook and Twitter status updates. J. Med. Internet Res. 20 (5), e168. https://doi.org/10.2196/jmir.9267.

Sekulić, I., Gjurković, M., Šnajder, J., 2019. Not Just Depressed: Bipolar Disorder Prediction on Reddit. ArXiv, 1811.04655 [Cs]. http://arxiv.org/abs/1811.04655.

Shen, J.H., Rudzicz, F., 2017. Detecting Anxiety Through Reddit. Proceedings of the Fourth Workshop on Computational Linguistics and Clinical Psychology—From Linguistic Signal to Clinical Reality, pp. 58–65. https://doi.org/10.18653/v1/W17-3107.

Tana, J.C., Kettunen, J., Eirola, E., Paakkonen, H., 2018. Diurnal variations of depression-related health information seeking: case study in Finland using Google Trends data. JMIR Ment. Health 5 (2), e43. https://doi.org/10.2196/mental.9152.

The Office of the National Coordinator for Health Information Technology, 2017. Health IT Quick Stats. Health IT Dashboard quickstats.

Tran, T., Kavuluru, R., 2017. Predicting mental conditions based on "history of present illness" in psychiatric notes with deep neural networks. J. Biomed. Inf. 75, S138–S148. https://doi.org/10.1016/j.jbi.2017.06.010.

Vermani, M., Marcus, M., Katzman, M.A., 2011. Rates of detection of mood and anxiety disorders in primary care: a descriptive, cross-sectional study. Prim. Care Compan. CNS Disord. 13 (2). https://doi.org/10.4088/PCC.10m01013.

Wang, R., Chen, F., Chen, Z., Li, T., Harari, G., Tignor, S., Zhou, X., Ben-Zeev, D., Campbell, A.T., 2014. Studentlife: Assessing Mental Health, Academic Performance and Behavioral Trends of College Students Using Smartphones. Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing—UbiComp '14 Adjunct, pp. 3–14. https://doi.org/10.1145/2632048.2632054.

Wang, R., Aung, M.S.H., Abdullah, S., Brian, R., Campbell, A.T., Choudhury, T., Hauser, M., Kane, J., Merrill, M., Scherer, E.A., Tseng, V.W.S., Ben-Zeev, D., 2016. CrossCheck: Toward Passive Sensing and Detection of Mental Health Changes in People With Schizophrenia. Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, pp. 886–897. https://doi.org/10.1145/2971648.2971740.

Wang, R., Wang, W., Obuchi, M., Scherer, E., Brian, R., Ben-Zeev, D., Choudhury, T., Kane, J., Hauser, M., Walsh, M., Campbell, A., 2020. On Predicting Relapse in Schizophrenia Using Mobile Sensing in a Randomized Control Trial, 2020 IEEE International Conference on Pervasive Computing and Communications (PerCom), pp. 1–8. https://doi.org/10.1109/PerCom45495.2020.9127365.

Ware, S., Yue, C., Morillo, R., Lu, J., Shang, C., Kamath, J., Bamis, A., Bi, J., Russell, A., Wang, B., 2018. Large-Scale Automatic Depression Screening Using Meta-Data from WiFi Infrastructure. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2, pp. 195:1–195:27. https://doi.org/10.1145/3287073.

Whiteford, H.A., Degenhardt, L., Rehm, J., Baxter, A.J., Ferrari, A.J., Erskine, H.E., Charlson, F.J., Norman, R.E., Flaxman, A.D., Johns, N., Burstein, R., Murray, C.J., Vos, T., 2013. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. Lancet 382 (9904), 1575–1586. https://doi.org/10.1016/S0140-6736(13)61611-6.

Whiting, A., Williams, D., 2013. Why people use social media: a uses and gratifications approach. Qual. Mark. Res. Int. J. 16 (4), 362–369. https://doi.org/10.1108/QMR-06-2013-0041.

Williams, K.L., Galliher, R.V., 2006. Predicting depression and self-esteem from social connectedness, support, and competence. J. Soc. Clin. Psychol. 25 (8), 855–874. https://doi.org/10.1521/jscp.2006.25.8.855.

Yang, A.C., Tsai, S.J., Huang, N.E., Peng, C.K., 2011. Association of Internet search trends with suicide death in Taipei City, Taiwan, 2004–2009. J. Affect. Disord. 132 (1), 179–184. https://doi.org/10.1016/j.jad.2011.01.019.

Yue, C., Ware, S., Morillo, R., Lu, J., Shang, C., Bi, J., Kamath, J., Russell, A., Bamis, A., Wang, B., 2017. Fusing Location Data for Depression Prediction, 2017 IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced Trusted Computed, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), pp. 1–8. https://doi.org/10.1109/UIC-ATC.2017.8397515.

Zhu, C., Li, B., Li, A., Zhu, T., 2016. Predicting Depression from Internet Behaviors by Time-Frequency Features, 2016 IEEE/WIC/ACM International Conference on Web Intelligence (WI), pp. 383–390. https://doi.org/10.1109/WI.2016.0060.