

Original Paper

The Relationship Between Smartphone-Recorded Environmental Audio and Symptomatology of Anxiety and Depression: Exploratory Study

Daniel Di Matteo¹, MASc; Kathryn Fotinos², BSc (Hons); Sachintha Lokuge², BSc (Hons); Julia Yu², MSc; Tia Sternat^{2,3}, MPsy; Martin A Katzman^{2,3,4,5}, MD, FRCPC; Jonathan Rose¹, PhD, FRSC

¹The Centre for Automation of Medicine, The Edward S Rogers Sr Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada

²Stress Trauma Anxiety Rehabilitation Treatment Clinic for Mood and Anxiety Disorders, Toronto, ON, Canada

³Department of Psychology, Adler Graduate Professional School, Toronto, ON, Canada

⁴Department of Psychology, Lakehead University, Thunder Bay, ON, Canada

⁵The Northern Ontario School of Medicine, Thunder Bay, ON, Canada

Corresponding Author:

Daniel Di Matteo, MASc

The Centre for Automation of Medicine

The Edward S Rogers Sr Department of Electrical and Computer Engineering

University of Toronto

DL Pratt Building

6 King's College Road

Toronto, ON, M5S 3H5

Canada

Phone: 1 416 978 6992

Email: dandm@ece.utoronto.ca

Abstract

Background: *Objective* and *continuous* severity measures of anxiety and depression are highly valuable and would have many applications in psychiatry and psychology. A collective source of data for objective measures are the sensors in a person's smartphone, and a particularly rich source is the microphone that can be used to sample the audio environment. This may give broad insight into activity, sleep, and social interaction, which may be associated with quality of life and severity of anxiety and depression.

Objective: This study aimed to explore the properties of passively recorded environmental audio from a subject's smartphone to find potential correlates of symptom severity of social anxiety disorder, generalized anxiety disorder, depression, and general impairment.

Methods: An Android app was designed, together with a centralized server system, to collect periodic measurements of the volume of sounds in the environment and to detect the presence or absence of English-speaking voices. Subjects were recruited into a 2-week observational study during which the app was run on their personal smartphone to collect audio data. Subjects also completed self-report severity measures of social anxiety disorder, generalized anxiety disorder, depression, and functional impairment. Participants were 112 Canadian adults from a nonclinical population. High-level features were extracted from the environmental audio of 84 participants with sufficient data, and correlations were measured between the 4 audio features and the 4 self-report measures.

Results: The regularity in daily patterns of activity and inactivity inferred from the environmental audio volume was correlated with the severity of depression ($r=-0.37$; $P<.001$). A measure of sleep disturbance inferred from the environmental audio volume was also correlated with the severity of depression ($r=0.23$; $P=.03$). A proxy measure of social interaction based on the detection of speaking voices in the environmental audio was correlated with depression ($r=-0.37$; $P<.001$) and functional impairment ($r=-0.29$; $P=.01$). None of the 4 environmental audio-based features tested showed significant correlations with the measures of generalized anxiety or social anxiety.

Conclusions: In this study group, the environmental audio was shown to contain signals that were associated with the severity of depression and functional impairment. Associations with the severity of social anxiety disorder and generalized anxiety disorder were much weaker in comparison and not statistically significant at the 5% significance level. This work also confirmed previous work showing that the presence of voices is associated with depression. Furthermore, this study suggests that sparsely sampled audio volume could provide potentially relevant insight into subjects' mental health.

(*JMIR Form Res* 2020;4(8):e18751) doi: [10.2196/18751](https://doi.org/10.2196/18751)

KEYWORDS

depression; anxiety; mobile phone; ecological momentary assessment; mobile apps; mobile health; digital signal processing; acoustics; speech recognition software

Introduction

Background

Depression and anxiety disorders are some of the most prevalent mental health disorders [1], yet access to services and treatment for these disorders is lacking. It is common for many Ontario residents with mental health problems to wait for 6 months to 1 year for treatment [2]. Automation of a part of the mental health care process may help address this service gap in our health care system, as automated assessment could alleviate some of the workload that is currently being carried by health care workers.

The health care process can be modeled as beginning with assessment and measurement, followed by diagnosis, and finally treatment. Subsequent rounds of measurement or assessment occur with the final goal of achieving remission. This work focuses on the measurement and diagnosis components by working toward building an automated and objective severity measurement of anxiety, depression, and functional impairment associated with poor mental health.

Research in both psychiatry and clinical psychology traditionally involves assessments of subjects' state (eg, mood and behavior) in clinical or research settings where they are removed from their natural home and living environment. Often, these assessments were performed retrospectively, where the subjects were asked to recollect behaviors and feelings over several weeks in the past. Ecological momentary assessment (EMA) [3] is an alternative approach that endeavors to assess subjects' mood and behaviors in a naturalistic setting. It occurs in real time and has a higher frequency of measurement, possibly multiple times a day. These assessments are frequent and occur in subjects' natural setting(s), removing clinicians (and the potential for bias), from the measurement process. Furthermore, EMA is not limited to self-report data but can also use data collected from sensors (physiological sensors and smartphone sensors). Sensor-based data are especially interesting in an EMA context because they can be collected *passively* without any interaction from the study subject, essentially addressing concerns of self-report biases surrounding self-reported data [4,5]. Researchers now use the term *passive EMA* [6] to refer to EMA systems in which sensor-based data are collected without any interaction from the user. It is also sometimes referred to as *unobtrusive EMA* or *passive sensing*.

The feasibility of building passive EMA systems has been greatly improved by the smartphone revolution. Smartphones

are ubiquitous and affordable consumer electronics, which are equipped with a wide range of sensors that can enable the type of sensing or monitoring necessary to perform passive EMA [7]. A brief survey of some of the existing work using passive EMA in the mental health space follows, including a more detailed review of studies that have used ecological audio data to predict mental health state.

Previous Work

A general methodology in these passive EMA or mobile sensing studies, and one that is used in this work, is to compute metrics, or *features*, from objective data sources that are designed to capture behaviors or traits of subjects that are known or suspected to be predictive of mental state. These features condense a large number of data points from a data source (eg, thousands of GPS coordinates measured over weeks of a study) into a single metric of behavior. This metric of behavior can then be tested for correlation with clinical measures of subjects' mental state. One such example of a GPS location-derived feature is the proportion of time a subject spends outside home. A low proportion might be indicative of avoidance behavior or low energy and, therefore, relevant to depression, for example.

A systematic review by Rohani et al [8] examined correlations between passively sensed smartphone data and symptoms of depression. *Homestay*, the proportion of time spent by the subject at home (computed from GPS data), and *screen active duration*, the proportion of time spent using the phone were 2 of the most strongly correlated features with depression. These features were reported as significant by numerous studies included in the review [8].

This general methodology of sampling objective data from subjects' smartphones (or other digital sensors) to infer health characteristics has been used in numerous studies, across many conditions. Although we will provide a focused review of relevant works that have used audio data to predict or measure mood and anxiety disorders, there is a wealth of research that has looked at using many different data sources to investigate, predict, or measure the severity of many characteristics of health and mental health disorders. Interested readers are directed to work that has investigated subjects' general mood and mental health [9-15], substance abuse [16,17], depression [18-24], bipolar disorder [25-29], anxiety disorders [30-32], and schizophrenia [33,34]. The most commonly used sources of passively collected smartphone data in these works include subjects' geolocation (ie, GPS data), screen activity and phone

usage time, SMS and phone metadata, and physical activity and motion sensor data.

As all smartphones are equipped with microphones, they can be used to detect audio-based features of a subject's environment. Several works have investigated the recording and analysis of speech audio from subjects' smartphones. There are different strategies to record audio, ranging from (1) actively prompting users to speak into a microphone, (2) passively recording subjects' phone calls, and (3) to passively recording environmental audio with no interaction from the user.

Using the active prompt-style methodology, Dickerson et al [35] conducted a study of depression in which subjects were asked twice daily to respond verbally to a prompt in free-form speech recorded by a microphone, yielding responses that were, on average, 1 to 2 min long. These audio recordings of prompted speech were then analyzed to produce 2 features: the fundamental frequency of subjects' speaking voices (F_0) and subjects' speech pause time. These 2 features were used to build a linear model for predicting the mood of the subjects. Mood was measured on a 1 to 10 scale (on the continuum of extremely depressed to extremely elevated mood), and the linear model was able to predict mood scores with a residual error of 0.092 (12 degrees of freedom, $P=.011$) [35]. Similarly, Guidi et al [36] investigated the fundamental frequency of speech (F_0) actively recorded from subjects via prompts in a study of bipolar disorder. This study was able to distinguish between individual bipolar subject's mood states. A total of 7 features extracted from speech audio were subjected to Kruskal-Wallis tests, and all features showed significant differences (at a 5% significance level) across angry, neutral, bored, and happy emotional states.

Using the passive phone call recording-style methodology, Faurholt-Jepsen et al [29] conducted a study of 28 outpatient subjects with bipolar disorder. Voice features produced from patients' phone calls were used to build 2 classification models that classified patients' states as manic or mixed versus euthymic (area under the curve=0.89) and depressive versus euthymic (area under the curve=0.78). Another study of bipolar subjects conducted by Grünerbl et al [25] used features of subjects' voices produced from the recordings of phone calls to predict mood state with 70% accuracy.

Finally, audio can be sampled in a much more passive and pervasive manner by using a smartphone's microphone to record environmental (ambient) audio. The StudentLife study by Wang et al [37] sampled ambient audio and used audio analysis techniques to detect the presence of human voices in the environment as a proxy measure of conversation frequency. They found that conversation frequency has a significant negative correlation with self-reported measures of depression severity [37]. The work by Abdullah et al [26] used the same approach in a study of 7 subjects with bipolar disorder. The conversation frequency feature was found to be weakly correlated with mood patterns ($r=0.16$; $P=.06$) as measured by the social rhythm metric. Ben-Zeev et al [10] used the amount of time proximal to human speech as a predictive feature in a study of general mental health. In a functional regression analysis, the amount of time spent proximal to human voices was found to be significantly associated with changes in a

self-reported measure of depressive symptoms over the course of the study ($P=.048$).

Goal of This Study

This exploratory study seeks to discover potential correlates of anxiety and depression symptomatology from environmental audio acquired using passive smartphone sensing. Although previous research has studied how some features of the audio environment relate to depression and bipolar mood disorders, we will extend this to include anxiety disorders. In addition, one aspect of our study is the exploration of the sampled *average volume of the environment* over time, which has privacy-preserving attributes. We are not aware of any other study of mental health that makes use of this measurement. We hypothesize that the time series of the volume of subjects' environments reveals some qualities and characteristics of their daily activities, which are associated with symptoms of depression and anxiety, a hypothesis that we believe has not been investigated in the literature. In addition, we explore the effect of the presence of voices in the sampled audio.

Methods

Overview

Subjects from a nonclinical population were recruited for a 2-week observational study in which a custom app was installed on their personal Android phone. Self-report measures of anxiety, depression, and general quality of life and impairment were collected at the beginning and end of the study. Throughout the duration of the study, the smartphone app passively collected the average volume of environmental audio and the presence of voice activity (whether or not speech was detected in the environment at the time of recording). A set of features was designed and used to extract higher-level information from this set of data, and a statistical analysis was performed to determine if a significant relationship existed between subjects' self-reported anxiety, depression, and general impairment and these features. The study was approved by the University of Toronto Health Sciences Research Ethics Board (Protocol #36687).

Recruitment

Subjects were recruited from Prolific [38], a web-based platform for recruiting study participants. Prolific maintains an active pool of subjects who wish to engage in research activities and enables researchers to deploy web-based tasks to subjects with specified demographics. It is similar to other services, such as Amazon's Mechanical Turk [39], but has some properties that make it more attractive to academic researchers. These include ethical payment requirements and a comprehensive database of subjects' demographic data to enable prescreening.

The study inclusion criteria were as follows: subjects should (1) reside in Canada, (2) be fluent in English, (3) own an Android phone, (4) have completed at least 95% of their previous Prolific studies successfully, and (5) have previously participated in at least 20 Prolific studies. The final criterion was used to ensure that subjects were proficient at using the Prolific system and were generally technology literate. There

were no exclusion criteria for the study. Subjects were paid Can \$18.50 (US \$14) for participating in the study.

Study Procedure

Members of the Prolific community who met the inclusion criteria could read a description of the study, which included an informed consent guide. Those who consented to the study were then directed to a webpage that acted as the study entry point. This website directed subjects to install the app from the Google Play app store and provided them with log-in credentials for using the study app. Once installed, the study app guided subjects through a short setup, where they were asked to provide the app with the necessary permissions to access their data, followed by a log-in. Immediately following setup and log-in, subjects were asked to complete a set of 4 self-report measures in digital form within the study app. At this point, following the completion of the self-report measures, the app began to periodically collect data in the background. No further actions or interactions with the study app were performed until the end of the study, exactly 14 days later, at the same time of day as the app installation/self-report work. At this time, subjects received a notification on their phone, informing them that the study had ended and requesting that they complete the same set of 4 self-report measures done at the beginning, again in the smartphone app. Following completion of this task, subjects were directed to uninstall the app from their phone and mark their study tasks as complete on the Prolific website. Subjects were then paid through Prolific's payment system.

Self-Report Measures

Subjects completed 4 self-report measures in digital form within the study app at the beginning and end of the 14-day study. A review by Belisario et al [40] found that self-administered survey scores do not differ when deployed by app versus other delivery modes. The 4 measures were the Liebowitz Social Anxiety Scale (LSAS) [41], the 7-item Generalized Anxiety Disorder Scale (GAD-7) [42], the 8-item Patient Health Questionnaire Scale (PHQ-8) [43], and the Sheehan Disability Scale (SDS) [44]. The LSAS is a 24-item scale used to assess the symptoms of social anxiety disorder by measuring respondents' fear and avoidance of various social situations [41]. The LSAS was originally developed as a clinician-administered instrument, yet the self-report version has been shown to have good psychometric properties [45]. The GAD-7 is a 7-item self-report scale used as a screener and severity measure of generalized anxiety disorder [42]. The PHQ-8 is an abbreviated form of the Patient Health Questionnaire 9-item depression scale [46], which omits the final item of the PHQ-9, a question that assesses suicidal ideation. The PHQ-8, similar to the PHQ-9, has been shown to be a valid diagnostic and severity measure for depressive disorders [43]. The PHQ-8 was chosen instead of the PHQ-9 because, owing to the anonymous and remote nature of the data collection in this study, the study investigators would be unable to properly intervene in the case of any evidence of risk for self-harm. Finally, the SDS is a 3-item self-report measure of general impairment, which has been shown to be a sensitive tool for measuring mental health-related functional impairment [44].

Both the GAD-7 and PHQ-8 instruments ask subjects to evaluate their symptoms over the past 2 weeks, whereas the LSAS and SDS ask subjects to evaluate their symptoms over the past week. Therefore, 2 weeks was the shortest duration possible to encompass the largest time window of assessment of the self-report measures, which is the rationale behind a 2-week study duration.

Smartphone Data Collection

An Android app was designed and created to collect all study data. This includes both the self-reported measures, described earlier, and the passively collected audio data—the volume of environmental audio and the presence or absence of speaking voices in the environment.

The study app records audio every 5 min, for a duration of 15 seconds, by turning on the microphone and recording the environment. This recording process occurs without any interaction from the user and with no notification to the user. Audio recordings are then securely transmitted from subjects' smartphones over the internet to a computer server where 2 further processing steps are performed. First, the average volume of each 15-second audio recording was calculated using the FFmpeg audio processing software framework [47]. Second, the presence of voices in the audio recording was detected using the Google Cloud Speech-to-Text software product [48]. This software generates text transcripts from audio recordings of speech; it was used to detect the presence of speech in audio recordings by simply noting whether each audio recording generated a transcript. Audio files containing silence, noise, or unintelligible speech do not successfully generate a text transcript, whereas recordings that contain intelligible speech produce a transcript.

The audio sampling period was chosen to be 5 min as a good trade-off between large amounts of data (with a shorter period) and the preservation of battery life of subjects' smartphones (with a longer period). Internal testing before the study showed a 5-min sampling period to be satisfactory for preserving battery life. Although a shorter period could yield more data, versions of the Android operating system since version 6 prevent this. Specifically, devices are prevented from performing background processing (such as this type of microphone sampling) while the device is sleeping more than once in a 9-min period [49]. Thus, for many devices, we are already at the limit of how frequently we can sample data in the background.

Data Preprocessing

Preprocessing of the volume time series was performed before feature extraction to account for missing data and to perform normalization. Periodic audio recordings were not reliably produced at a precise period of 5 min by the study app, so the volume time series were resampled to a period of 5 min, and missing samples were imputed using linear interpolation. After resampling and interpolation, volume samples were clipped at the ceiling and floor of 3 SDs from the mean of the volume time series to remove outliers (using subject mean and subject SD, not group). Finally, the volume time series were scaled linearly to ensure that all volume measurements were within the range

of 0 to 1. No preprocessing of the voice presence time series was performed.

Feature Extraction

This subsection describes the methods used to compute the 4 correlates of anxiety and depression symptomatology derived from subjects' environmental audio recordings. These correlates, or *features*, were extracted from the volume and speech presence time series data to test for the association with symptoms of anxiety and depression as measured by the LSAS, GAD-7, PHQ-8, and SDS. In the sections below, we describe 3 features from the volume time series, called *daily similarity*, *sleep disturbance-all nights*, and *sleep disturbance-weeknights*. A fourth feature was extracted from the speech presence time series called the *speech presence ratio*.

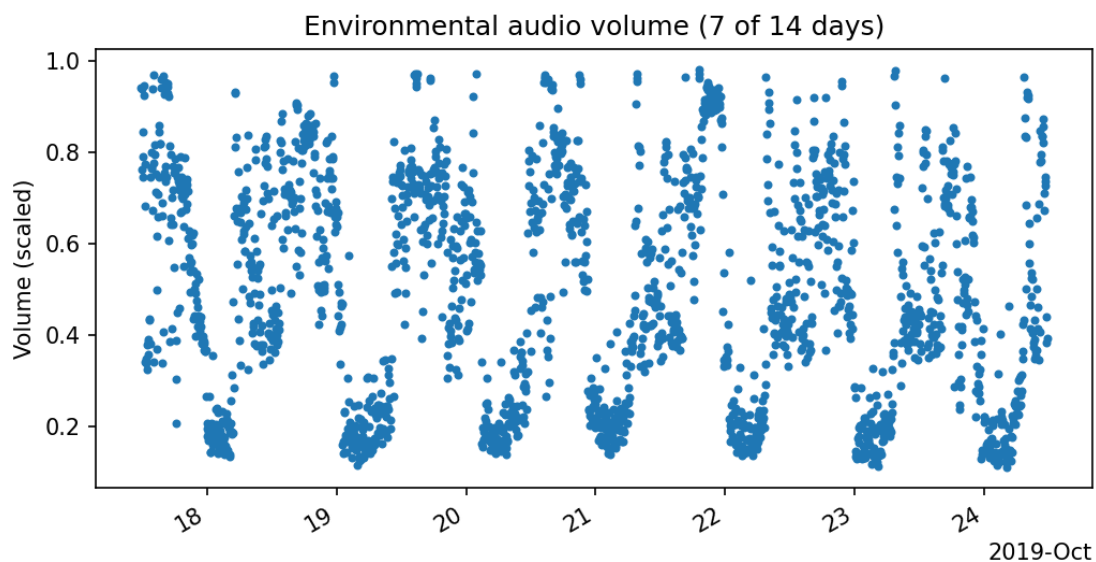
Daily Similarity

The daily similarity feature was designed to infer the consistency of the subjects' sequence of daily activities. Visualizations of the volume time series clearly show distinct periods of activity (characterized by large spikes in volume) and inactivity

(characterized by quieter volume with less variance). These periods coincide roughly with daytime and nighttime, respectively. Furthermore, these patterns are periodic and repeat daily. Figure 1 shows a visualization of 7 days of a subject's environmental audio volume data.

A link between regularity in daily activities, including sleep, and anxiety and depression is commonly described in the literature [50,51], and therefore, it follows that this feature may be associated with symptoms of anxiety and depression. To quantify the regularity of this pattern of daily activity, the *autocorrelation* function was computed for each subject's volume time series. This is a signal processing technique that computes the correlation between a signal and a time-delayed copy of itself [52]. The autocorrelation function of a signal is a time-dependent Pearson correlation coefficient of the signal and its copy for varying degrees of time lag between the two. As we are interested in quantifying the similarity between a subject's days, the value of the autocorrelation function at a time lag of 24 hours is most relevant. The *daily similarity* feature is, therefore, defined as the value of the autocorrelation function of the volume time series evaluated at a lag time of 24 hours.

Figure 1. Visualization of a subject's environmental audio volume data (7 of 14 days).



Sleep Disturbance

Majority of the periods of apparent inactivity that are visible in the volume time series appear to coincide with sleep. It was hypothesized that a proxy measure of sleep quality can be inferred by measuring how chaotic the volume of subjects' environments are during sleep times, replicating the link between sleep and mental health reported in the literature. For example, the prevalence of mood disorders has been shown to be much higher in populations with chronic sleep problems [53], and insomnia may be a state marker of anxiety disorders [54]. It follows, therefore, that a feature that infers the quality of subjects' sleep may be associated with symptoms of anxiety and depression.

To quantify sleep quality, the volume time series was examined with periods of quiet noted to be characterized by low variance in volume. Although the absolute value of the volume is also low at quiet times, the threshold for what can be considered

quiet is greatly dependent on the specific microphone and phone placement; therefore, variance was considered a more appropriate measure of the noisiness of the environment. The *sleep disturbance* feature is defined as the SD of the volume time series between the hours of 12:00 AM and 06:00 AM, local time. This is a proxy measure of sleep quality. More specifically, it is a measure of the noisiness of a subject's environment during common hours of sleep. In addition, as people's patterns of sleep can vary between weeknights and weekends, we compute 2 versions of this feature: (1) using volume data between the hours of 12:00 AM and 06:00 AM on all nights of the week, and (2) using only volume data from weeknights (Monday through Friday). These choices are made without specific knowledge of the subjects' workday or workweek schedule.

Speech Presence Ratio

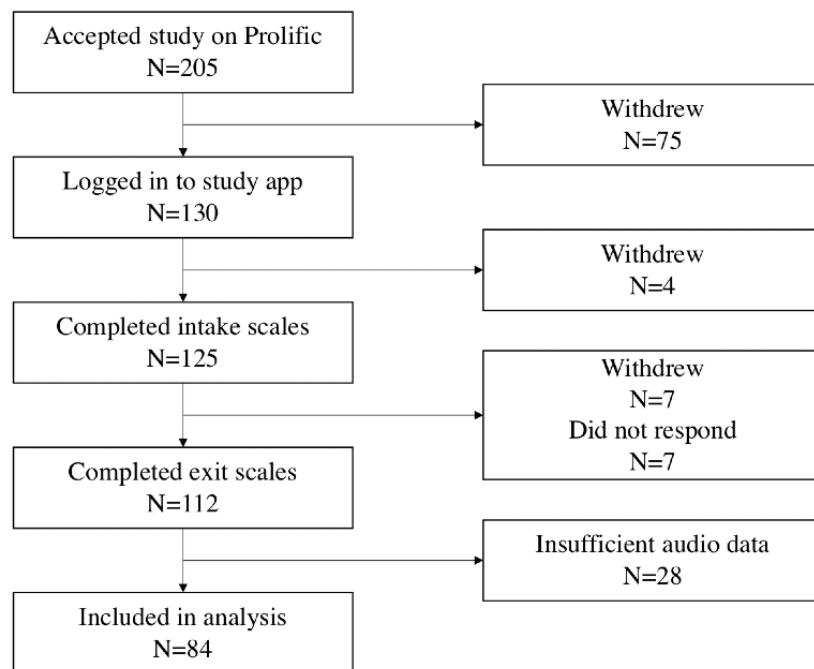
Previous studies of mental health using mobile sensing have computed proxy measures of social interaction as a feature

predictive of depression severity [10,37]. To replicate these results and to ascertain if this also holds for measures of social anxiety, generalized anxiety, and general impairment, a proxy measure of social interaction was computed. The *speech presence ratio* feature is computed from the speech presence time series data and is defined as the proportion of audio recordings containing the presence of speech to the total number of audio recordings.

Privacy

A number of privacy considerations drove the design of the study procedure, app, and data collection. Prolific, the platform from which subjects were recruited, anonymizes subjects. Subjects were provided with app log-in credentials, which were provided on demand to each subject as they enrolled in the study to avoid subjects using their name, email address, or some other potentially identifying information as their log-in name. Audio recordings were encrypted both at rest (on subjects' phones) and in transit to the server. Once processed on the server side, audio files were deleted. The speech transcripts generated to detect the presence of speech were processed in the following way: each transcript was broken into pairs of words (ie, bigrams) and then stored in random order for use in later studies. The stripping of ordering and time information from bigrams was done to prevent later re-creation of transcripts.

Figure 2. Flow chart of study recruitment.



Study Group Characteristics

The study sample had an average age of 30 years (SD 8.6) and 42% (35/84) of subjects were female. The mean and SD of the 4 self-reported measures are presented in Table 1. To understand whether the self-reported measures may be related to the age and sex of subjects, 2 statistics were computed. The Pearson correlation coefficient was computed to measure the correlation between self-reported measures and subject age, and an independent samples *t* test was performed to test whether the mean scale score differed between the male and female sexes.

Results

Overview and Data Inclusion

From July 2019 to December 2019, 205 eligible Prolific members entered the study. Withdrawals were common, with 86 subjects choosing to withdraw at some point in the study (commentary on the high withdrawal rate is provided in the Limitations subsection of the Discussion section). Of the 119 subjects who did not withdraw, 112 completed both sets of self-report questionnaires. Finally, 84 of the 112 completed subjects yielded sufficient audio data for analysis based on the criterion that at least 50% of the expected number of audio recordings were made by the study app. Figure 2 provides an illustration of the study recruitment results. Although self-report measures were collected at both the beginning and end of the 2-week study, all results that followed used the values collected at the end of the study. The postobservation values of the self-report measures were used in the analysis because if such a smartphone-based assessment system were to exist, it would be more useful to predict trajectories or upcoming symptom severity, not severity 2 weeks before the start of data recording (as would be the case if correlations/associations with the prestudy scales were measured).

The LSAS, GAD-7, and SDS were negatively correlated with subject age, but there was no significant difference in the mean scale scores of the 2 sexes (at a 5% significance level). Table 1 also lists the results of these tests.

To further characterize the mental health of our study subjects, self-report measures were used to screen for social anxiety disorder, generalized anxiety disorder, and major depressive disorder. A cutoff of 60 was used with the LSAS scores to screen for social anxiety disorder (generalized subtype), as recommended by Mennin et al [55]. A cutoff of 10 was used

with the GAD-7 scores to screen for generalized anxiety disorder. This cutoff was shown to optimize sensitivity (89%) and specificity (82%) [42]. A cutoff of 10 was used with the PHQ-8 scores to screen for depression, as recommended by Kroenke et al [43]. Table 2 summarizes the results of these

screenings on our study sample. The rates of social anxiety (32%), generalized anxiety (26%), and depression (37%) in our study sample were all compared with the rates in the general Canadian population. This will be elaborated in the Discussion section.

Table 1. Descriptive statistics for self-report measures of anxiety and depression (n=84).

Measures	Score, mean (SD)	Correlation with age		Difference between mean scores of the sexes	
		<i>r</i>	<i>P</i> value	<i>t</i> test (<i>df</i>)	<i>P</i> value
Liebowitz Social Anxiety Scale	53.7 (25.8)	-0.27	.01	-1.68 (82)	.10
Generalized Anxiety Disorder seven-item scale	6.6 (4.6)	-0.29	.01	-1.37 (82)	.17
Patient Health Questionnaire eight-item scale	8.5 (5.6)	-0.19	.09	-1.18 (82)	.24
Sheehan Disability Scale	10.8 (7.7)	-0.26	.02	-1.12 (82)	.27

Table 2. Results of screening the study sample for depression and anxiety disorders (n=84).

Disorders	Screening criteria	Positive screenings, n (%)
Social anxiety disorder	Liebowitz Social Anxiety Scale score ≥ 60	32 (38)
Generalized anxiety disorder	Generalized Anxiety Disorder seven-item scale score ≥ 10	22 (26)
Major depressive disorder	Patient Health Questionnaire eight-item scale score ≥ 10	31 (37)

Objective Audio Features

The objective audio features described in subsection *Feature Extraction* of the Methods section were computed using the environmental audio time series data of the 84 study subjects. Descriptive statistics for the 4 features are presented in Table 3. These descriptive statistics include the mean and SD of the

features observed in the study sample, and additionally the minimum value observed, the maximum value observed, and the values of the first three quartiles. The values of the speech presence ratio appear to align reasonably with intuition, with subjects spending anywhere from 1% to 30% of their time in the presence of speech.

Table 3. Descriptive statistics for objective audio features (n=84).

Features	Mean (SD)	Minimum	Q1	Q2	Q3	Maximum
Daily similarity	0.80 (0.07)	0.45	0.77	0.83	0.85	0.90
Sleep disturbance—all nights	0.14 (0.06)	0.03	0.10	0.13	0.17	0.32
Sleep disturbance—weeknights	0.14 (0.06)	0.03	0.09	0.12	0.18	0.34
Speech presence ratio	0.15 (0.06)	0.01	0.11	0.16	0.20	0.30

Correlations Between Audio Features and Self-Report Measures

To test the association between the audio features and the self-reported measures of anxiety, depression, and functional impairment, the Pearson correlation coefficient was computed between each feature and scale. The daily similarity feature is negatively correlated with all 4 self-report measures, which supports the hypothesis that regularity in daily activity and circadian rhythm is associated with more positive mental health (ie, lower scale scores). This feature was most strongly correlated with depressive symptoms. The sleep disturbance

feature, whether computed using all nighttime audio or only weeknight audio, was positively correlated with all 4 self-report measures, which is in line with the hypothesis that better sleep quality (ie, less sleep disturbance) is associated with positive mental health. The strength of the correlation is improved when only the weeknight audio is considered. Finally, the speech presence ratio feature was negatively correlated with all 4 self-report measures, where the correlation with depressive symptoms was the strongest for all observed correlations ($r=-0.37$; $P<.001$). Table 4 summarizes the results of the correlation analysis. No correction for multiple tests was performed, as we view this work as exploratory.

Table 4. Pearson correlation between objective audio features and self-reported measures of anxiety and depression (n=84).

Features	Liebowitz Social Anxiety Scale		Generalized Anxiety Disorder seven-item scale		Patient Health Questionnaire eight-item scale		Sheehan Disability Scale	
	<i>r</i>	<i>P</i> value	<i>r</i>	<i>P</i> value	<i>r</i>	<i>P</i> value	<i>r</i>	<i>P</i> value
Daily similarity	-0.20	.07	-0.19	.09	-0.37	<.001	-0.18	.10
Sleep disturbance—all nights	0.00	.99	0.07	.52	0.17	.13	0.15	.17
Sleep disturbance—weeknights	0.05	.65	0.12	.26	0.23	.03	0.18	.11
Speech presence ratio	-0.19	.08	-0.16	.14	-0.37	<.001	-0.29	.01

Discussion

Population

The self-report measures completed by the subjects revealed that this study's sample, despite being recruited from a healthy population, had a high prevalence of depression and anxiety. Data reported by the Government of Canada in 2006 estimate a 12-month prevalence of major depressive disorder at 4.8% [1], compared with a positive screening rate of 37% in our sample. The same 2006 report lists a 12-year prevalence of the combined class of anxiety disorders at 4.8%, a figure that is significantly lower than the positive screening rates for generalized anxiety disorder (26%) and social anxiety disorder (38%) observed in our sample. Given that the cutpoints used for screening were all shown to have high specificity, it seems unlikely that these high rates are solely a result of false-positive screenings. Instead, 2 possible explanations are proposed, which may be jointly responsible. First, the Canadian population of subjects on the Prolific recruitment platform may have elevated rates of mood and anxiety disorders with respect to the general Canadian population. It may be that people who choose to find work on the platform go there because of these conditions, which might prevent them from doing outside-the-home work. Second, that subject sampling was impacted by self-selection bias. In other words, Prolific participants who struggle with mental health to some degree may be more likely to have chosen to enroll and remain in a study that focuses on mental health.

Negative correlations were measured between age and self-reported measures of social anxiety ($r=-0.27$; $P=.01$), generalized anxiety ($r=-0.29$; $P=.01$), depression ($r=-0.19$; $P=.09$), and general impairment owing to poor mental health ($r=-0.26$; $P=.02$) in our sample of subjects. This indicates that younger subjects generally displayed worse mental health than older subjects. It is not clear if this observation is part of a more general trend in the greater population or if it is specific to the Prolific population. A review of studies examining the occurrence of anxiety, depression, and distress found some evidence that aging is associated with less susceptibility to anxiety and depression, but it is unclear if that was because of aging or cohort effects [56]. A study of mood disorders and suicide using data on American adolescents and adults from 2005 to 2017 observed an increase in mood disorders and suicidal ideation since the mid-2000s, and it is suggested that cultural trends in the use of digital media among younger individuals may be responsible for creating a cohort effect [57].

Correlations

The key finding of this work is the development and evaluation of a set of features, computed from subjects' environmental audio, as potential correlates of symptoms of anxiety, depression, and functional impairment. Correlation analysis of these features and self-reported measures, summarized in Table 4, shows that all 4 features are associated more strongly with symptoms of depression (measured by the PHQ-8) than with any other symptoms. The daily similarity and speech presence ratio were both significantly correlated with PHQ-8 scores at a 5% significance level. Speech presence ratio, but not daily similarity, was likewise significantly correlated with SDS scores at a 5% significance level. None of the 4 features were found to be significantly correlated with either LSAS or GAD-7 scores at a 5% significance level.

We note that while associations with the LSAS and GAD-7 are weak or nonexistent, the associations with the SDS are nearly as strong as those with the PHQ-8. This may suggest that the impairment that is being measured by the SDS may, in large part, be due to symptoms of depression. Indeed, we measured a stronger correlation between the SDS and the PHQ-8 scores of subjects ($r=0.76$; $P<.001$) than between the SDS and the LSAS ($r=0.48$; $P<.001$) or the GAD-7 ($r=0.62$; $P<.001$).

The fact that some features are associated with depressive symptomatology but none are associated with the symptomatology of generalized anxiety or social anxiety disorder is interesting, and we offer some speculation as to why this may be the case. First, we must observe that our features are very coarse—they measure sleep, activity, and speech, but with no specific context (they are measured on a gross scale). Depression is broadly debilitating on energy and activity, and if this impact is independent of a specific context, our features are appropriately designed to detect this impact. This is in contrast to anxiety disorders, which often have specific triggers that are context dependent. Individuals with anxiety can avoid contexts that trigger their anxiety and, therefore, present as if they do not suffer from the effects of anxiety as long as they continue to exhibit avoidance behaviors. Although it is unlikely that anxious individuals are able to avoid all triggers, especially those with generalized subtypes, avoidance behavior may be partially responsible for weakening the associations between inferred behavior and mental state. To passively measure the severity of anxiety disorders, it seems that any feature must capture avoidance behavior as a proxy for anxiety itself and

also have some measure of state anxiety to detect when a subject is in a specific context that acts as an anxiety trigger.

Comparison With Other Studies

To our knowledge, no other studies have inferred daily patterns of activity and inactivity solely from volume samples of ambient audio, which is captured by the daily similarity feature, so direct comparisons of the daily similarity feature with other known works are not possible. However, a feature called circadian movement, first proposed by Saeb et al [20], captures the degree to which the pattern of a subject's visits to different geolocations follows a 24-hour or circadian rhythm. The study by Saeb et al [20] measured a significant negative correlation between circadian movement and PHQ-9 scores ($r=-0.63$; $P=.01$; $n=18$). Our ambient volume-based measure of daily regularity in activity, called daily similarity, was also negatively correlated with PHQ-8 scores ($r=-0.37$; $P<.001$; $n=84$). This further supports the hypothesis that regularity in daily activities is associated with lower severity of depressive symptoms. It is interesting to consider that the same underlying signal may be present in both GPS data and ambient audio data. A study by Ware et al [58] also used the circadian movement feature in the prediction of depression. Although this feature was one of many used in a classifier that achieved an F1 score as high as 0.86, it is unclear to what degree this feature alone was associated with the severity of depressive symptoms.

The association between sleep and mental health has been investigated in a number of studies. Self-reported measures of sleep quality have been shown to be associated with state anxiety [59], anxiety disorders [60], and depression [61,62]. Sleep duration has been measured objectively in mobile sensing studies and shown to have a significant association with the severity of depressive symptoms [10,23,37]. Although sleep duration and sleep disturbance features are different measures and as such cannot be directly compared, these studies support our results in demonstrating a general association between sleep and depression.

Finally, although the speech presence ratio feature does not appear in identical form in the literature, there are other studies that have used similar proxy measures of social interaction as correlates of depression severity. The pioneering study by Wang et al [37] inferred the number of conversations that subjects encountered throughout their days by performing an analysis of ambient audio and also found a significant negative correlation with PHQ-9 scores ($r=-0.39$; $P=.02$; $n=48$). This specific feature by Wang et al is more contextually aware than our speech presence ratio because not all audio that they found to contain speech was considered conversational in their system. Specifically, they ignored speech if it was detected during lecture or group meeting hours of their student subject population. Despite the fact that our proxy measure of social interaction is context unaware, we also show a negative correlation between the speech presence ratio and PHQ-8 scores ($r=-0.37$; $P<.001$; $n=84$). Time spent proximal to human speech was also found to be significantly associated with changes in PHQ-9 scores ($P=.048$) in a study by Ben-Zeev et al [10].

Limitations

A fundamental limitation of the study design is that as a cross-sectional study, it is not possible to make any claim regarding causation between the observed features and severity of anxiety, depression, or functional impairment. It cannot be determined, for example, if avoiding social contact (as inferred by the speech presence ratio feature) causes increased depression severity or if an increase in depression severity owing to some other factor causes individuals to retreat socially and engage in less speaking.

A high proportion of subject withdrawals can also be considered as a limitation of this study (86/205 subjects, 42%, chose to withdraw). The majority of withdrawals, 87% (75/86), occurred before a successful log-in to the study app (see Figure 2 for an illustration of the withdrawals throughout the study timeline). Having observed how most withdrawals occurred early in the study, we provide 2 possible hypotheses for the high number of withdrawals. The first is the relative difficulty in setting up the study app on a personal smartphone, which is more complex than the typical task asked of subjects recruited on the Prolific platform. Our setup procedure included turning off the smartphone's battery optimizations for the study app, which requires some facility with Android settings. The second is the possibility that subjects are unwilling to provide the app with the permissions necessary for data collection. The study app asks users to grant permissions before log-in, which might explain why 87% (75/86) of the withdrawals occurred without logging in. Our previous study explored clinical patients' willingness to consent to the collection of different forms of data collection for mental health purposes [54]. The number of withdrawals is in line with the data from the previous study.

A further limitation surrounding subject withdrawals is the possibility of sample bias. The 86 individuals who withdrew from the study may differ from those who remained in the study. We are unable to test this because Prolific removes researchers' access to the demographic data (age and sex) of participants who withdraw from studies. Nondemographic data (ie, digital data and self-report measures) may also differ between the groups, yet this is difficult to test because 92% (79/86) of withdrawals withdrew early enough in the study so as not to provide any digital data or self-report measures. Of the 112 individuals who remained in the study and completed all tasks, it is possible to test for differences between the group included in the analysis (ie, the group with sufficient audio data) and the group excluded from the analysis. These 2 groups did not differ significantly in age or on any of the 4 self-report measures, either at intake or exit, as tested by *t* tests at a 5% significance level (two-sided). These 2 groups also did not differ by sex (counts of men and women), as tested by a chi-square test (1 degree of freedom) for independence at a 5% significance level.

Some limitations also exist regarding the validity of the features. The speech presence ratio feature does not distinguish between recorded speech (eg, from a TV or radio) and human speech, it simply detects intelligible speech. This method does not distinguish between speakers, so in many cases, the subject themselves may not be the person speaking. The method also only detects English speech, so it will potentially miss speech

if, for example, a subject does not speak English at home. Finally, our technique for detecting speech using automatic speech recognition is more biased toward false negatives than false positives. If speech is detected by the system, it is highly likely that the speech is present, yet it is much more likely to miss speech in noisy environments or environments with multiple speakers speaking concurrently.

The sleep disturbance feature is affected by the subjects' specific mobile phone hardware, where different microphones with different automatic gain control functionality (which dynamically adjusts the volume and is not controllable by programmers) could produce different measurements in the same environment. To produce perfectly consistent features, one would be required to use a device such as a calibrated sound level meter, which measures volume as an absolute measure with no gain control.

Finally, it must be noted that the feasibility of completely passive mobile sensing with a high frequency of data sampling is becoming increasingly difficult on Android devices. Battery optimization features limit the rate at which apps can turn on

in the background and sample their environment [63] because any app activity that *wakes up* a device not actively being used will drain device battery. Although Google does offer mechanisms for limiting battery optimizations to improve the frequency at which passive sensing can occur, many Android device vendors add third-party battery optimization software that can be difficult to disable systematically. This vendor-specific constraint on apps is a significant issue because of the fragmented nature of the Android ecosystem.

Conclusions

This work contributes toward the development of automated and objective severity measurements of anxiety, depression, and functional impairment associated with poor mental health. Focusing solely on environmental audio, which was passively sensed from subjects' smartphones, this work presents 2 new correlates of depressive symptoms and general impairment, which we refer to as the daily similarity and sleep disturbance features. Furthermore, this work supports previous findings by reproducing a measured association between time spent proximal to speech and severity of depression.

Acknowledgments

This research was funded in part by the Natural Science and Engineering Research Council Discovery Grant number RGPIN-2019-04395.

Conflicts of Interest

MAK has been a consultant or advisory board member for GlaxoSmithKline, Lundbeck, Eli Lilly, Boehringer Ingelheim, Organon, AstraZeneca, Janssen, Janssen-Ortho, Solvay, Bristol-Myers Squibb, Shire, Sunovion, Pfizer, Purdue, Merck, Astellas, Tilray, Bedrocan, Takeda, Eisai, and Otsuka. MAK conducted research for GlaxoSmithKline, Lundbeck, Eli Lilly, Organon, AstraZeneca, Janssen-Ortho, Solvay, Genuine Health, Shire, Bristol-Myers Squibb, Takeda, Pfizer, Hoffman La Roche, Biotics, Purdue, Astellas, Forest, and Lundbeck. MK has received honoraria from GlaxoSmithKline, Lundbeck, Eli Lilly, Boehringer Ingelheim, Organon, AstraZeneca, Janssen, Janssen-Ortho, Solvay, Bristol-Myers Squibb, Shire, Sunovion, Pfizer, Purdue, Merck, Astellas, Bedrocan, Tilray, Allergan, and Otsuka. MAK has received research grants from the Canadian Institutes of Health Research, Sick Kids Foundation, Centre for Addiction and Mental Health Foundation, Canadian Psychiatric Research Foundation, Canadian Foundation for Innovation, and the Lotte and John Hecht Memorial Foundation.

References

1. The Human Face of Mental Health and Mental Illness in Canada. Public Health Agency of Canada (PHAC). 2006. URL: https://www.phac-aspc.gc.ca/publicat/human-humain06/pdf/human_face_e.pdf [accessed 2020-02-18]
2. Annual Report. Office of the Auditor General of Ontario. 2016. URL: https://www.auditor.on.ca/en/content/annualreports/arreports/en16/2016AR_v1_en_web.pdf [accessed 2020-02-18]
3. Moskowitz DS, Young SN. Ecological momentary assessment: what it is and why it is a method of the future in clinical psychopharmacology. *J Psychiatry Neurosci* 2006 Jan;31(1):13-20 [FREE Full text] [Medline: 16496031]
4. van de Mortal TH. Faking it: social desirability response bias in self-report research. *Aust J Adv Nurs* 2008;25(4):40-48 [FREE Full text]
5. John OP, Robins RW. Accuracy and bias in self-perception: individual differences in self-enhancement and the role of narcissism. *J Pers Soc Psychol* 1994 Jan;66(1):206-219. [doi: 10.1037//0022-3514.66.1.206] [Medline: 8126650]
6. Ruwaard J, Kooistra L, Thong M. Ecological Momentary Assessment in Mental Health Research: A Practical Introduction With Examples in R. EMA Research Manual. 2018. URL: https://jruwaard.github.io/aph_ema_handbook/ [accessed 2020-02-18]
7. Vaid S, Harari G. Smartphones in personal informatics: a framework for self-tracking research with mobile sensing. In: *Digital Phenotyping and Mobile Sensing: New Developments in Psychoinformatics*. New York, USA: Springer; 2019.
8. Rohani DA, Faurholt-Jepsen M, Kessing LV, Bardram JE. Correlations between objective behavioral features collected from mobile and wearable devices and depressive mood symptoms in patients with affective disorders: systematic review. *JMIR Mhealth Uhealth* 2018 Aug 13;6(8):e165 [FREE Full text] [doi: 10.2196/mhealth.9691] [Medline: 30104184]
9. Sandstrom GM, Lathia N, Mascolo C, Rentfrow PJ. Putting mood in context: using smartphones to examine how people feel in different locations. *J Res Personal* 2017 Aug;69:96-101. [doi: 10.1016/j.jrp.2016.06.004]

10. Ben-Zeev D, Scherer EA, Wang R, Xie H, Campbell AT. Next-generation psychiatric assessment: using smartphone sensors to monitor behavior and mental health. *Psychiatr Rehabil J* 2015 Sep;38(3):218-226 [FREE Full text] [doi: [10.1037/prj0000130](https://doi.org/10.1037/prj0000130)] [Medline: [25844912](https://pubmed.ncbi.nlm.nih.gov/25844912/)]
11. Asselbergs J, Ruwaard J, Ejdys M, Schrader N, Sijbrandij M, Riper H. Mobile phone-based unobtrusive ecological momentary assessment of day-to-day mood: an explorative study. *J Med Internet Res* 2016 Mar 29;18(3):e72 [FREE Full text] [doi: [10.2196/jmir.5505](https://doi.org/10.2196/jmir.5505)] [Medline: [27025287](https://pubmed.ncbi.nlm.nih.gov/27025287/)]
12. Pratap A, Atkins DC, Renn BN, Tanana MJ, Mooney SD, Anguera JA, et al. The accuracy of passive phone sensors in predicting daily mood. *Depress Anxiety* 2019 Jan;36(1):72-81. [doi: [10.1002/da.22822](https://doi.org/10.1002/da.22822)] [Medline: [30129691](https://pubmed.ncbi.nlm.nih.gov/30129691/)]
13. Baras K, Soares L, Paulo N, Barros R. 'Smartphone': Supporting Students' Well-Being According to Their Calendar and Mood. In: International Multidisciplinary Conference on Computer and Energy Science. 2016 Presented at: SpliTech'16; July 13-15, 2016; Split, Croatia. [doi: [10.1109/splitech.2016.7555919](https://doi.org/10.1109/splitech.2016.7555919)]
14. Mark G, Czerwinski M, Iqbal S, Johns P. Workplace Indicators of Mood: Behavioral and Cognitive Correlates of Mood Among Information Workers. In: Proceedings of the 6th International Conference on Digital Health Conference. 2016 Presented at: DH'16; April 11-13, 2016; Montreal, Canada. [doi: [10.1145/2896338.2896360](https://doi.org/10.1145/2896338.2896360)]
15. Cho YM, Lim HJ, Jang H, Kim K, Choi JW, Shin C, et al. A cross-sectional study of the association between mobile phone use and symptoms of ill health. *Environ Health Toxicol* 2016;31:e2016022 [FREE Full text] [doi: [10.5620/eh.t.2016022](https://doi.org/10.5620/eh.t.2016022)] [Medline: [27788568](https://pubmed.ncbi.nlm.nih.gov/27788568/)]
16. Lee H, Ahn H, Choi S, Choi W. The SAMS: smartphone addiction management system and verification. *J Med Syst* 2014 Jan;38(1):1. [doi: [10.1007/s10916-013-0001-1](https://doi.org/10.1007/s10916-013-0001-1)] [Medline: [24395031](https://pubmed.ncbi.nlm.nih.gov/24395031/)]
17. Naughton F, Hopewell S, Lathia N, Schallbroeck R, Brown C, Mascolo C, et al. A context-sensing mobile phone app (Q sense) for smoking cessation: a mixed-methods study. *JMIR Mhealth Uhealth* 2016 Sep 16;4(3):e106 [FREE Full text] [doi: [10.2196/mhealth.5787](https://doi.org/10.2196/mhealth.5787)] [Medline: [27637405](https://pubmed.ncbi.nlm.nih.gov/27637405/)]
18. Burns MN, Begale M, Duffecy J, Gergle D, Karr CJ, Giangrande E, et al. Harnessing context sensing to develop a mobile intervention for depression. *J Med Internet Res* 2011 Aug 12;13(3):e55 [FREE Full text] [doi: [10.2196/jmir.1838](https://doi.org/10.2196/jmir.1838)] [Medline: [21840837](https://pubmed.ncbi.nlm.nih.gov/21840837/)]
19. Saeb S, Lattie EG, Schueller SM, Kording KP, Mohr DC. The relationship between mobile phone location sensor data and depressive symptom severity. *PeerJ* 2016;4:e2537 [FREE Full text] [doi: [10.7717/peerj.2537](https://doi.org/10.7717/peerj.2537)] [Medline: [28344895](https://pubmed.ncbi.nlm.nih.gov/28344895/)]
20. Saeb S, Zhang M, Karr CJ, Schueller SM, Corden ME, Kording KP, et al. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. *J Med Internet Res* 2015 Jul 15;17(7):e175 [FREE Full text] [doi: [10.2196/jmir.4273](https://doi.org/10.2196/jmir.4273)] [Medline: [26180009](https://pubmed.ncbi.nlm.nih.gov/26180009/)]
21. Wahle F, Kowatsch T, Fleisch E, Rufer M, Weidt S. Mobile sensing and support for people with depression: a pilot trial in the wild. *JMIR Mhealth Uhealth* 2016 Sep 21;4(3):e111 [FREE Full text] [doi: [10.2196/mhealth.5960](https://doi.org/10.2196/mhealth.5960)] [Medline: [27655245](https://pubmed.ncbi.nlm.nih.gov/27655245/)]
22. Farhan AA, Yue C, Morillo R, Ware S, Lu J, Bi J, et al. Behavior vs Introspection: Refining Prediction of Clinical Depression via Smartphone Sensing Data. In: Wireless Health. 2016 Presented at: WH'16; October 25-27, 2016; Bethesda, MD, USA p. 1-8. [doi: [10.1109/wh.2016.7764553](https://doi.org/10.1109/wh.2016.7764553)]
23. Demasi O, Aguilera A, Recht B. Detecting Change in Depressive Symptoms From Daily Wellbeing Questions, Personality, and Activity. In: Wireless Health. 2016 Presented at: WH'16; October 25-27, 2016; Bethesda, MD p. 1-8. [doi: [10.1109/wh.2016.7764552](https://doi.org/10.1109/wh.2016.7764552)]
24. Canzian L, Musolesi M. Trajectories of Depression: Unobtrusive Monitoring of Depressive States by Means of Smartphone Mobility Traces Analysis. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2015 Presented at: UbiComp'15; September 7-11, 2015; Osaka, Japan. [doi: [10.1145/2750858.2805845](https://doi.org/10.1145/2750858.2805845)]
25. Grünerbl A, Muaremi A, Osmani V, Bahle G, Ohler S, Tröster G, et al. Smartphone-based recognition of states and state changes in bipolar disorder patients. *IEEE J Biomed Health Inform* 2015 Jan;19(1):140-148. [doi: [10.1109/JBHI.2014.2343154](https://doi.org/10.1109/JBHI.2014.2343154)] [Medline: [25073181](https://pubmed.ncbi.nlm.nih.gov/25073181/)]
26. Abdullah S, Matthews M, Frank E, Doherty G, Gay G, Choudhury T. Automatic detection of social rhythms in bipolar disorder. *J Am Med Inform Assoc* 2016 May;23(3):538-543. [doi: [10.1093/jamia/ocv200](https://doi.org/10.1093/jamia/ocv200)] [Medline: [26977102](https://pubmed.ncbi.nlm.nih.gov/26977102/)]
27. Beiwinkel T, Kindermann S, Maier A, Kerl C, Moock J, Barbian G, et al. Using smartphones to monitor bipolar disorder symptoms: a pilot study. *JMIR Ment Health* 2016 Jan 6;3(1):e2 [FREE Full text] [doi: [10.2196/mental.4560](https://doi.org/10.2196/mental.4560)] [Medline: [26740354](https://pubmed.ncbi.nlm.nih.gov/26740354/)]
28. Faurholt-Jepsen M, Vinberg M, Frost M, Debel S, Christensen EM, Bardram JE, et al. Behavioral activities collected through smartphones and the association with illness activity in bipolar disorder. *Int J Methods Psychiatr Res* 2016 Dec;25(4):309-323 [FREE Full text] [doi: [10.1002/mpr.1502](https://doi.org/10.1002/mpr.1502)] [Medline: [27038019](https://pubmed.ncbi.nlm.nih.gov/27038019/)]
29. Faurholt-Jepsen M, Busk J, Frost M, Vinberg M, Christensen EM, Winther O, et al. Voice analysis as an objective state marker in bipolar disorder. *Transl Psychiatry* 2016 Jul 19;6:e856 [FREE Full text] [doi: [10.1038/tp.2016.123](https://doi.org/10.1038/tp.2016.123)] [Medline: [27434490](https://pubmed.ncbi.nlm.nih.gov/27434490/)]
30. Boukhechba M, Chow P, Fua K, Teachman BA, Barnes LE. Predicting social anxiety from global positioning system traces of college students: feasibility study. *JMIR Ment Health* 2018 Jul 4;5(3):e10101 [FREE Full text] [doi: [10.2196/10101](https://doi.org/10.2196/10101)] [Medline: [29973337](https://pubmed.ncbi.nlm.nih.gov/29973337/)]

31. Chow PI, Fua K, Huang Y, Bonelli W, Xiong H, Barnes LE, et al. Using mobile sensing to test clinical models of depression, social anxiety, state affect, and social isolation among college students. *J Med Internet Res* 2017 Mar 3;19(3):e62 [FREE Full text] [doi: [10.2196/jmir.6820](https://doi.org/10.2196/jmir.6820)] [Medline: [28258049](https://pubmed.ncbi.nlm.nih.gov/28258049/)]
32. Boukhechba M, Daros AR, Fua K, Chow PI, Teachman BA, Barnes LE. DemonicSalmon: monitoring mental health and social interactions of college students using smartphones. *Smart Health* 2018 Dec;9-10:192-203. [doi: [10.1016/j.smhl.2018.07.005](https://doi.org/10.1016/j.smhl.2018.07.005)]
33. Ben-Zeev D, Wang R, Abdullah S, Brian R, Scherer EA, Mistler LA, et al. Mobile behavioral sensing for outpatients and inpatients with schizophrenia. *Psychiatr Serv* 2016 May 1;67(5):558-561 [FREE Full text] [doi: [10.1176/appi.ps.201500130](https://doi.org/10.1176/appi.ps.201500130)] [Medline: [26695497](https://pubmed.ncbi.nlm.nih.gov/26695497/)]
34. di Francesco S, Fraccaro P, van DV. Out-of-Home Activity Recognition from GPS Data in Schizophrenic Patients. In: 29th International Symposium on Computer-Based Medical Systems. 2016 Presented at: CBMS'16; June 20-24, 2016; Dublin, Ireland. [doi: [10.1109/cbms.2016.54](https://doi.org/10.1109/cbms.2016.54)]
35. Dickerson RF, Gorlin EI, Stankovic JA. Empath: A Continuous Remote Emotional Health Monitoring System for Depressive Illness. In: Proceedings of the 2nd Conference on Wireless Health. 2011 Presented at: WH'11; October 10-13, 2011; Atlanta, GA, USA. [doi: [10.1145/2077546.2077552](https://doi.org/10.1145/2077546.2077552)]
36. Guidi A, Vanello N, Bertschy G, Gentili C, Landini L, Scilingo E. Automatic analysis of speech F0 contour for the characterization of mood changes in bipolar patients. *Biomed Signal Process Control* 2015 Mar;17:29-37. [doi: [10.1016/j.bspc.2014.10.011](https://doi.org/10.1016/j.bspc.2014.10.011)]
37. Wang R, Chen F, Chen Z. StudentLife: Assessing Mental Health, Academic Performance and Behavioral Trends of College Students Using Smartphones. In: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2014 Presented at: UbiComp'14; September 7-11, 2014; Osaka, Japan. [doi: [10.1145/2632048.2632054](https://doi.org/10.1145/2632048.2632054)]
38. Prolific. URL: <https://www.prolific.co/> [accessed 2019-07-01]
39. Mortensen K, Hughes TL. Comparing Amazon's mechanical Turk platform to conventional data collection methods in the health and medical research literature. *J Gen Intern Med* 2018 Apr;33(4):533-538 [FREE Full text] [doi: [10.1007/s11606-017-4246-0](https://doi.org/10.1007/s11606-017-4246-0)] [Medline: [29302882](https://pubmed.ncbi.nlm.nih.gov/29302882/)]
40. Belisario JS, Jamsek J, Huckvale K, O'Donoghue J, Morrison CP, Car J. Comparison of self-administered survey questionnaire responses collected using mobile apps versus other methods. *Cochrane Database Syst Rev* 2015 Jul 27(7):MR000042. [doi: [10.1002/14651858.MR000042.pub2](https://doi.org/10.1002/14651858.MR000042.pub2)] [Medline: [26212714](https://pubmed.ncbi.nlm.nih.gov/26212714/)]
41. Liebowitz MR. Social phobia. *Mod Probl Pharmacopsychiatry* 1987;22:141-173. [doi: [10.1159/000414022](https://doi.org/10.1159/000414022)] [Medline: [2885745](https://pubmed.ncbi.nlm.nih.gov/2885745/)]
42. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006 May 22;166(10):1092-1097. [doi: [10.1001/archinte.166.10.1092](https://doi.org/10.1001/archinte.166.10.1092)] [Medline: [16717171](https://pubmed.ncbi.nlm.nih.gov/16717171/)]
43. Kroenke K, Strine TW, Spitzer RL, Williams JB, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. *J Affect Disord* 2009 Apr;114(1-3):163-173. [doi: [10.1016/j.jad.2008.06.026](https://doi.org/10.1016/j.jad.2008.06.026)] [Medline: [18752852](https://pubmed.ncbi.nlm.nih.gov/18752852/)]
44. Leon AC, Olfson M, Portera L, Farber L, Sheehan DV. Assessing psychiatric impairment in primary care with the Sheehan disability Scale. *Int J Psychiatry Med* 1997;27(2):93-105. [doi: [10.2190/T8EM-C8YH-373N-1UWD](https://doi.org/10.2190/T8EM-C8YH-373N-1UWD)] [Medline: [9565717](https://pubmed.ncbi.nlm.nih.gov/9565717/)]
45. Baker SL, Heinrichs N, Kim H, Hofmann SG. The liebowitz social anxiety scale as a self-report instrument: a preliminary psychometric analysis. *Behav Res Ther* 2002 Jun;40(6):701-715. [doi: [10.1016/s0005-7967\(01\)00060-2](https://doi.org/10.1016/s0005-7967(01)00060-2)] [Medline: [12051488](https://pubmed.ncbi.nlm.nih.gov/12051488/)]
46. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001 Sep;16(9):606-613 [FREE Full text] [doi: [10.1046/j.1525-1497.2001.016009606.x](https://doi.org/10.1046/j.1525-1497.2001.016009606.x)] [Medline: [11556941](https://pubmed.ncbi.nlm.nih.gov/11556941/)]
47. FFmpeg Developers. FFmpeg. 2019. URL: <http://ffmpeg.org/> [accessed 2020-02-07]
48. Speech-To-Text. Google Cloud Platform. URL: <https://cloud.google.com/speech-to-text/> [accessed 2020-02-07]
49. Optimize for Doze and App Standby. Android Developers. URL: <https://developer.android.com/training/monitoring-device-state/doze-standby> [accessed 2020-06-08]
50. Shear MK, Randall J, Monk TH, Ritenour A, Tu X, Frank E, et al. Social rhythm in anxiety disorder patients. *Anxiety* 1994;1(2):90-95. [doi: [10.1002/anxi.3070010208](https://doi.org/10.1002/anxi.3070010208)] [Medline: [9160553](https://pubmed.ncbi.nlm.nih.gov/9160553/)]
51. Szuba MP, Yager A, Guze BH, Allen EM, Baxter LR. Disruption of social circadian rhythms in major depression: a preliminary report. *Psychiatry Res* 1992 Jun;42(3):221-230. [doi: [10.1016/0165-1781\(92\)90114-j](https://doi.org/10.1016/0165-1781(92)90114-j)] [Medline: [1496054](https://pubmed.ncbi.nlm.nih.gov/1496054/)]
52. Gubner J. Probability and Random Processes for Electrical and Computer Engineers. Cambridge, UK: Cambridge University Press; 2006.
53. Benca RM, Okawa M, Uchiyama M, Ozaki S, Nakajima T, Shibui K, et al. Sleep and mood disorders. *Sleep Med Rev* 1997 Nov;1(1):45-56. [doi: [10.1016/s1087-0792\(97\)90005-8](https://doi.org/10.1016/s1087-0792(97)90005-8)] [Medline: [15310523](https://pubmed.ncbi.nlm.nih.gov/15310523/)]
54. Neckelmann D, Mykletun A, Dahl AA. Chronic insomnia as a risk factor for developing anxiety and depression. *Sleep* 2007 Jul;30(7):873-880 [FREE Full text] [doi: [10.1093/sleep/30.7.873](https://doi.org/10.1093/sleep/30.7.873)] [Medline: [17682658](https://pubmed.ncbi.nlm.nih.gov/17682658/)]
55. Mennin DS, Fresco DM, Heimberg RG, Schneier FR, Davies SO, Liebowitz MR. Screening for social anxiety disorder in the clinical setting: using the Liebowitz social anxiety scale. *J Anxiety Disord* 2002;16(6):661-673. [doi: [10.1016/s0887-6185\(02\)00134-2](https://doi.org/10.1016/s0887-6185(02)00134-2)] [Medline: [12405524](https://pubmed.ncbi.nlm.nih.gov/12405524/)]
56. Jorm AF. Does old age reduce the risk of anxiety and depression? A review of epidemiological studies across the adult life span. *Psychol Med* 2000 Jan;30(1):11-22. [doi: [10.1017/s0033291799001452](https://doi.org/10.1017/s0033291799001452)] [Medline: [10722172](https://pubmed.ncbi.nlm.nih.gov/10722172/)]

57. Twenge JM, Cooper AB, Joiner TE, Duffy ME, Binau SG. Age, period, and cohort trends in mood disorder indicators and suicide-related outcomes in a nationally representative dataset, 2005-2017. *J Abnorm Psychol* 2019 Apr;128(3):185-199. [doi: [10.1037/abn0000410](https://doi.org/10.1037/abn0000410)] [Medline: [30869927](https://pubmed.ncbi.nlm.nih.gov/30869927/)]
58. Ware S, Yue C, Morillo R, Lu J, Shang C, Bi J, et al. Predicting depressive symptoms using smartphone data. *Smart Health* 2020 Mar;15:100093. [doi: [10.1016/j.smhl.2019.100093](https://doi.org/10.1016/j.smhl.2019.100093)]
59. Atalay H. Comorbidity of insomnia detected by the Pittsburgh sleep quality index with anxiety, depression and personality disorders. *Isr J Psychiatry Relat Sci* 2011;48(1):54-59 [FREE Full text] [Medline: [21572444](https://pubmed.ncbi.nlm.nih.gov/21572444/)]
60. Ramsawh HJ, Stein MB, Belik S, Jacobi F, Sareen J. Relationship of anxiety disorders, sleep quality, and functional impairment in a community sample. *J Psychiatr Res* 2009 Jul;43(10):926-933. [doi: [10.1016/j.jpsychires.2009.01.009](https://doi.org/10.1016/j.jpsychires.2009.01.009)] [Medline: [19269650](https://pubmed.ncbi.nlm.nih.gov/19269650/)]
61. Hayashino Y, Yamazaki S, Takegami M, Nakayama T, Sokejima S, Fukuhara S. Association between number of comorbid conditions, depression, and sleep quality using the Pittsburgh sleep quality index: results from a population-based survey. *Sleep Med* 2010 Apr;11(4):366-371. [doi: [10.1016/j.sleep.2009.05.021](https://doi.org/10.1016/j.sleep.2009.05.021)] [Medline: [20219425](https://pubmed.ncbi.nlm.nih.gov/20219425/)]
62. Ağargün MY, Kara H, Solmaz M. Subjective sleep quality and suicidality in patients with major depression. *J Psychiatr Res* 1997;31(3):377-381. [doi: [10.1016/s0022-3956\(96\)00037-4](https://doi.org/10.1016/s0022-3956(96)00037-4)] [Medline: [9306295](https://pubmed.ncbi.nlm.nih.gov/9306295/)]
63. Power Management. Android Developers. URL: <https://developer.android.com/about/versions/pie/power> [accessed 2020-02-07]

Abbreviations

- EMA:** ecological momentary assessment
GAD-7: 7-item Generalized Anxiety Disorder Scale
LSAS: Liebowitz Social Anxiety Scale
PHQ-8: Patient Health Questionnaire eight-item scale
SDS: Sheehan Disability Scale

Edited by M Focsa; submitted 21.03.20; peer-reviewed by P Chow, B Teachman; comments to author 20.05.20; revised version received 17.06.20; accepted 07.07.20; published 13.08.20

Please cite as:

Di Matteo D, Fotinos K, Lokuge S, Yu J, Sternat T, Katzman MA, Rose J

The Relationship Between Smartphone-Recorded Environmental Audio and Symptomatology of Anxiety and Depression: Exploratory Study

JMIR Form Res 2020;4(8):e18751

URL: <https://formative.jmir.org/2020/8/e18751>

doi: [10.2196/18751](https://doi.org/10.2196/18751)

PMID: [32788153](https://pubmed.ncbi.nlm.nih.gov/32788153/)

©Daniel Di Matteo, Kathryn Fotinos, Sachintha Lokuge, Julia Yu, Tia Sternat, Martin A Katzman, Jonathan Rose. Originally published in JMIR Formative Research (<http://formative.jmir.org>), 13.08.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://formative.jmir.org>, as well as this copyright and license information must be included.