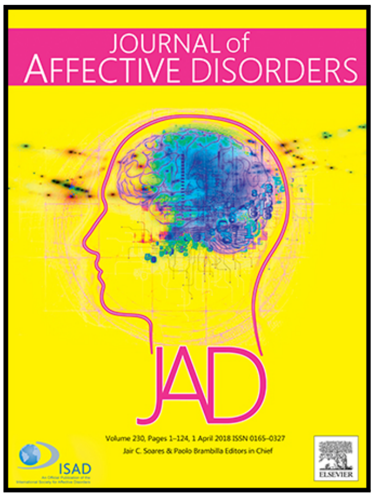
'This is the peer reviewed version of the following article: Knight, A., & Bidargaddi, N. (2018). Commonly available activity tracker apps and wearables as a mental health outcome indicator: A prospective observational cohort study among young adults with psychological distress. Journal of Affective Disorders, 236, 31–36. https://doi.org/10.1016/j.jad.2018.04.099

which has been published in final form at
https://doi.org/10.1016/j.jad.2018.04.099

Accepted Manuscript



Commonly available activity tracker apps and wearables as a mental health outcome indicator: A prospective observational cohort study among young adults with psychological distress

Alissa Knight , Niranjan Bidargaddi

PII: S0165-0327(17)32351-0
DOI: 10.1016/j.jad.2018.04.099
Reference: JAD 9754

To appear in: *Journal of Affective Disorders*

Received date: 15 November 2017
Revised date: 14 March 2018
Accepted date: 8 April 2018

Please cite this article as: Alissa Knight , Niranjan Bidargaddi , Commonly available activity tracker apps and wearables as a mental health outcome indicator: A prospective observational cohort study among young adults with psychological distress, *Journal of Affective Disorders* (2018), doi: 10.1016/j.jad.2018.04.099

Highlights

- Adaptive and dynamic just in-time monitoring using digital footprints was demonstrated.
- A high level of acceptance in sharing personalised data from activity trackers was found.
- Greater daily activity duration was detected from wearable devices over smartphone apps.
- Increased volatility was associated with high randomness and low predictability of physical activity in people with moderate levels of anxiety.

Commonly available activity tracker apps and wearables as a mental health outcome indicator: A prospective observational cohort study among young adults with psychological distress

Alissa Knight [a,1], Niranjan Bidargaddi [a,b]

[a] *Personal Health Informatics, College of Medicine and Public Health, Flinders University, Adelaide, Australia*
[b] *South Australian Health and Medical Research Institute, Adelaide, Australia*

[1] Corresponding Author: Dr Alissa Knight, Personal Health Informatics, Flinders University, Tonsley (Level 2), GPO Box 2100, Adelaide 5001, South Australia,
Email: alissa.knight@flinders.edu.au

Abstract

Background: Monitoring is integral to adequately recognise and track mental health indicators of symptoms and functioning. Early identification of warning signs from digital footprints could facilitate adaptive and dynamic just in-time monitoring and care for individuals with common mental disorders. **Methods:** Self-report data on mental health and lifestyle behaviour from 120 male and female Australian young adults experiencing psychological distress were collected online. API software was used to download participant's daily activity duration measurements over eight months from linked commercial activity tracker apps and wearables in real time. An independent samples *t*-test was conducted to compare the differences in daily durations of recorded physical activity between wearable devises and smartphone apps. Entropy techniques using R interpol package were used to analyse volatility in daily activity duration. **Results:** DASS-21 depression, stress and anxiety sub-scale scores indicated the study sample on average, had a moderate level of psychological distress. Daily activity duration was significantly greater from wearable devices when compared with smartphone apps (*t*-test = 25.4, p<0.001). Entropy indices were not related with any of the DASS-21 measures. However, significant correlation between

DASS-21 anxiety subscale scores and entropy of those with over 45 days measurements ($r = 0.58$, $p = 0.02$) was observed. **Limitations:** The observational nature of this study prohibits causal inference. As a convenience sample was used, the results may lack generalisability to the wider population. **Conclusions:** Continuous monitoring using commercial apps and wearables as a resource to help clinicians augment clinical care for common mental disorders appears viable.

Keywords: Common mental disorders; digital foot prints; daily activity duration; wearables; smartphone apps.

1. Introduction

Common mental disorders including depression, generalised anxiety disorder (GAD), panic disorder, phobias, social anxiety disorder, obsessive-compulsive disorder (OCD) and post-traumatic stress disorder (PTSD) affect approximately 8% of the world population (WHO, 2017). Historically, emphasis of burden has been placed on severe, organic mental disorders such as schizophrenia and bipolar disorder. In parallel to these types of neurological-based, chronic mental diseases, mental conditions considered to be more common in the community, while not typically considered 'severe', are by no means 'mild' (Goldberg, 1994). In fact, numerous studies have documented the immense suffering and disability associated with common mental disorders, as well as the broad, substantial social and economic detriment they generate. Indeed, a new study by the World Health Organization (WHO, 2016) estimated that the cost of depression and anxiety in over 36 countries around the world will amount to around $147 billion by 2030. Thus, there appears an urgent need to reshape the way common mental illnesses are foreseen and acted upon, including a multidisciplinary approach.

In 2002, the Australian Government attempted to address the impending mental illness epidemic with a new policy of ‘monitoring’ for all state-based community mental health services; instated in the Second National Mental Health Plan (DoHA, 2005). Indeed, monitoring is integral to adequately recognise and track mental health indicators of symptoms and functioning (Carlier, 2012). However, people with mental disorders are unprecedently difficult to monitor. For the most part, this is due to the fact that a substantial proportion of sufferers with common mental disorders do not often seek medical treatment (Pratt and Brody, 2014; Mojtabai et al., 2002). Hence, attempting to monitor such people in a way that facilitates the opportunity for early intervention, is particularly challenging (Collin et al., 2014).

Digital-based, continuous monitoring, has recently been documented as a responsive strategy that holds potential to target individuals suffering with common mental disorders in the community. Specifically, new research has shown that early identification of warning signs from digital footprints (i.e. passive data traces arising as a by-product of individuals' day-to-day interactions with mobile and/or Internet-connected technologies) could facilitate adaptive and dynamic just in-time monitoring and care for individuals with psychological distress (Bidargaddi et al., 2016).

Fixed-time monitoring systems, which are currently used in Australian public mental health services, are limited in their capacity to achieve this goal. In this approach, functional and symptomatic assessments are conducted only at fixed time points using traditional self-report and expert rated assessment measures. Typically, these approaches lack ecological validity as they are subject to recall biases, can only be administered at limited time points in controlled settings, and often require individuals to travel to medical settings to receive the assessment (Trull and Ebner-Priemer, 2012). This inability to assess the impact of environmental interactions with mental state in real-time, has hindered progress towards understanding and classifying mental disorders, as well as treating them. A more nuanced method for monitoring and detecting remote behavioural changes, which transports traditional clinic-based assessment into the naturalistic environments of individuals, may provide a better received therapeutic context that leads to early intervention and improved health outcomes.

Sensor signals automatically captured in smartphones and wearables may serve as an avenue to achieving such holistic view of patients' mental health trajectories in real-time (Patel et al., 2012). Indeed, there is much scope for incorporating consumer technologies into mental healthcare. Today, the vast majority of people today have "smartphones," a new advanced class of mobile phones. In fact, over 60% of the population worldwide currently own a mobile phone (4.61 billion in 2016), including among lower-socioeconomic groups and minority populations (Watson et al., 2016). Smartphones are a global positioning system (GPS), equipped with advanced computing and sensing capabilities, and host a number of sophisticated programs that can easily be installed on the phones (i.e., “apps”). These features, together with the unobtrusiveness of sensor data gathering and demonstrated evidence of popularity and acceptability, represents a potential information resource for epidemiological-based research in common mental disorders (Bidargaddi et al., 2016).

To date, the primary approach used by studies involves providing dedicated mobile phones to participants and developing custom apps to extract raw sensor signals (Capela et al., 2015). The advantage of this approach is better reliability and quality of data (del Rosario et al., 2015). However using dedicated devices has limited translational potential due to scalability issues arising from costs associated with providing devices. Introducing a new technology is also likely to alter participants default behaviour and environment, and thus the measurements might not reflect truest sense of the observant. An alternative approach is to derive indicators from sensors contained in smartphone apps or wearables that participants may already be using in their daily lives (Grunerbl et al., 2015). Collecting data via the use of smartphone apps in this way may reduce the likelihood of cofounders and/or bias typically associated with traditional forms of measuring behaviour. Furthermore, economically, the use of individually owned smartphones offers the potential of conducting assessment and providing intervention at immense scalability worldwide at little to no cost.

Activity tracker apps and wearables in particular, are uniquely positioned to capture the granular and temporal behavioural characteristics in natural settings, and reveal underlying objective indicators of a person's moment-to-moment functioning Grunerbl et al., 2015; Osmani, 2015). Indeed, studies have begun to show the efficacy of this capability among serious and severe mental illness populations. Grunerbl and colleagues (2015) were among the first to investigate whether physical activity, smartphone-sensing accelerometers could detect state changes (i.e. depressive and manic states) that could signal the onset of an episode in patients suffering from bipolar disorder. Results showed a strong correlation between physical activity daily interval (morning, afternoon, evening, night) scores and patients’ states (r=0.63, p<.05). Furthermore, within-patient Naïve Bayes classification results revealed 81% mean accuracy in recognising patients’ state, and 82% recall. Further recent studies have begun to emerge, supporting these findings for serious and severe mental illness populations (Grunerbl et al., 2015; Alvarez-Lozano et al., 2014; Mayora et al., 2013; Beiwinkel et al., 2016).

However, to date, scarce evidence from research supporting the use of activity tracker apps and wearables as a valid mental health assessment tool for common mental disorder populations exists. Therefore, the primary aim of the current study was to elicit a descriptive overview of young adults with common mental disorders using and sharing physical activity tracker app and wearable device data. Characteristics of this data, including processes of collection, ways to reconcile measurements from different activity tracker sources, missing data patterns, and approaches of deriving indicators for common mental disorder risk identification and potential interventional pathways are explored and discussed.

2. METHOD

2.1 Study design and recruitment

This study utilised a prospective observational cohort trial design over eight months conducted in Australia. A total of 120 male and female young Australian adults aged between 18-25 years old were initially recruited between July 2016 - September 2016 from the mental health web site ReachOut.com; an online mental health support platform targeted at Australian young adults between 13-25 (Vogl et al., 2016). A rolling recruitment strategy was used over a course of three months, involving two approaches: 1. emailing participants directly who had registered interest with ReachOut.com due to reported depressive, anxious or stressful psychological distress to participate in the study, and 2. placing popup banners about the study at the bottom right corner of some sections of the ReachOut website. Participation in the study was completely voluntary, and participants were informed at the beginning of the study they were free to withdraw at any stage.

2.2 Participant screening, selection and study procedure

The initial recruited sample of 120 volunteers who expressed interest in the study by either responding to email invitations or study advertisements were sent a detailed information pack about the study. Eligibility criteria for participation were: 1. aged between 18-25 years old, 2. currently using or had willingness to install activity tracker apps on a smartphone or wearable device, and 3. provided informed consent.

From an initial sample of 120 volunteers who were assessed for eligibility, 67 were excluded due to not meeting one or more of the outlined eligibility criteria, leaving 53 potential participants. As such, those 53 participants completed a detailed online survey about their health and lifestyle profile, provided authorised access to measurements recorded in their activity trackers apps through online API software, of which enabled data to be analysed. A flow diagram of participant's progress throughout the duration of the study is illustrated below in Figure 1.

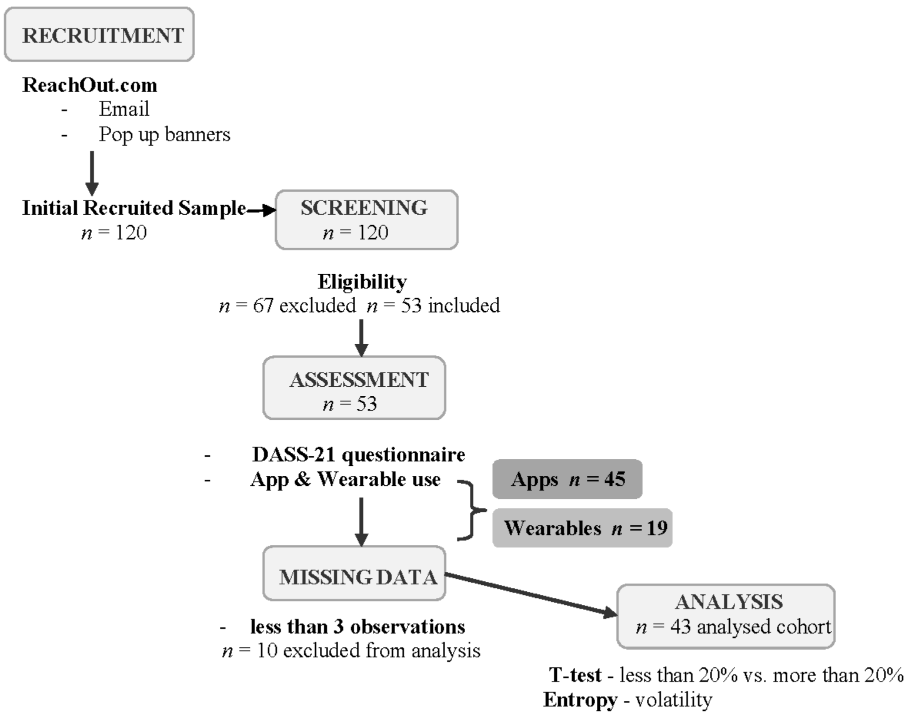


Figure 1. Flow diagram of study design

2.3 *Measures*

Common mental disorder symptomatology

The Depression Anxiety Stress Scale - 21 items (DASS-21) (Lovibond and Lovibond, 1995) was used in the current study as a quantitative measure of psychological distress along three axes of common mental disorder symptomatology; 1. Depression, 2. Anxiety, and 3. Stress. The DASS-21 *Depression* emotional state scale measures self-reported motivation, mood and self-esteem, while the DASS-21 *Anxiety* scale measures self-reported physiological arousal, perceived panic, and fear, and the DASS-21 *Stress* scale measures tension and irritability (Parkitny and McCauley, 2010). Each scale has seven individual items comprised of a series of statements and four short response options that reflect severity on a scale of 0 (did not apply to me at all) to 3 (applied to me very much, or most of the time). Lovebird and Lovebird (1995) provide normative values within the DASS manual that consider for *Depression*, a score of 0 – 9 as normal, 10 – 13 as mild, 14 – 20 as moderate, 21 – 27 as severe, and 28+ extremely severe; for *Anxiety*, a score of 0 – 7 as normal, 8 – 9 as mild, 10 – 14 as moderate, 15 – 19 as severe, and 20+ extremely severe; for *Stress*, a score of 0 – 14 as normal, 15 – 18 as mild, 19 – 25 as moderate, 26 – 33 as severe, and 34+ extremely severe. Internal consistency reliability for each of the subscales of the DASS-21 have been shown to be high (Cronbach’s alpha of 0.96 to 0.97 for the depression subscale, 0.84 to 0.92 for the anxiety subscale, and 0.90 to 0.95 for the stress subscale) (Parkitny and McCauley, 2010).

Daily activity duration

In this study, *daily activity duration* derived from commercially available activity tracker apps that participants were already using on their smartphones and wearables in their daily life was chosen as the primary outcome of interest. More specifically, micro-electro-mechanical system (MEMS) miniature motion sensors (i.e. location-based accelerometers) within smartphone apps were used to identify real-time, physical activity signals among participants. The accumulated information was transferred to a cloud API service (i.e. third-party proprietary system), encompassing a pre-processing data and feature extraction process, of which reduces the raw sensor data to a predetermined set of derived parameters. The web-based API software used in the study supported extraction of measurements from 26 different commercial activity tracker apps and/or wearable providers. The API software downloaded daily activity duration measurements from linked app in real time and stored them into a database. For each enrolled participant, daily activity duration was recorded for eight months post commencement of the study. In addition, corresponding dates of measurement and names of the measurement source were downloaded from the database. This approach facilitated the ability to infer the basic level of physical movement among participants.

3. Analysis

Statistical analysis was performed using SPSS Version 22 for Mac (Chicago, IL, USA), with an alpha set at $p< 0.05$ (two tailed). With respect to baseline characteristics, descriptive statistics were performed for categorical variables (frequencies and percentages), and for continuous variables the mean and standard deviation (SD) were calculated. All variables were screened for normality by way of visual inspection of the distribution in SPSS output, and dealt with according to standard methodology to obtain a normal distribution (Field, 2009). Differences in DASS-21 scores between the various group categories were analysed using t-tests: Group category 1. Those who used and shared daily activity duration measurements versus those who did not; Group category 2. App based versus device-based activity duration measurements; Group category 3. participants with greater than 20% missing activity duration measurements versus those with less than 20% missing activity duration. To analyse how the spread of missing data was over time, a continuous time series that included ‘no data’ for missing dates between recorded measurement dates was created for each participant. Subsequently, a plot of distribution of missing data over time was produced by aligning first measurement dates for all participants.

Patterns of volatility in daily activity duration observations were investigated through entropy analysis of participants who had less than 20% of missing observations. Missing daily activity duration observations between first and last available measurement date were imputed using five techniques (i.e. smooth average, random forest, mean substitution, interpolation, and omission) from R Interpol package. Secondly, activity duration measurements were normalised to create a discretized time series of activity duration by assigning each observed duration a value between zero to twenty-four after normalization. Third, entropy function in R package ‘entropy’ was applied on discretized activity duration time series to obtain a score in logarithmic scale.

Specifically, volatility in daily activity duration observations were calculated using the following formula:

$$H(X) = -\sum_i P(x_i) \log_{10} P(x_i),$$

Where *P* (xi) is the number of hours a participant was active on day xi. *P* (xi) can have discrete values between 0 (not active) to twelve (active for 12 hours or more). For example, in a mock-up scenario where a participant undergoes two hours of physical activity every day over 14 days (a total of 28 hours over 14 days), the resulting entropy would be:

```
y1 = c(2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2)
entropy.empirical(y1,unit="log10")
e1 = -sum(2/28 * log10(2/28)*14) = 1.15
```

Another scenario maybe where a participant is physically active for two hours over twelve days (weekdays + Sundays) and four hours on two days (Saturdays); a total of 32 hours over 14 days. The resulting entropy would be:

```
y2 = c(2, 2, 2, 2, 2, 4, 2, 2, 2, 2, 2, 2, 4, 2)
e2 = -(sum(2/32 * log10(2/32)*12) + sum(4/32 * log10(4/32)*2))
entropy.empirical(y2,unit="log10") = 1.13
```

Higher entropy in the first scenario is indicative of the fact that the participant (y1) spent time more uniformly across different days, while lower entropy in the second scenario reflects the participant (y2) had greater inequality in the time spent across days. As the time spend over time is less random, or less volatile, reflective of habits and routines; entropy scores go down.

Finally, relationship between entropy score and participant' DASS-21 characteristics were assessed using Pearson's correlation coefficient. Correlations were also investigated in sub-groups: participant's with under "45 days" and those over '45" observations, with 45 days chosen as the cut-off to reflect the typical intervention period of six weeks in mental health (NZGG, 2008).

4. Results

4.1 Descriptive Statistics

Participant Characteristics

Table 1. shows the means and standard deviations for the analysed cohort of 53 participants for the present analysis. Descriptive results showed that at baseline the mean age (±SD) of participants was 20.7 (3.2) years old. The vast majority of the study sample were female (77%), with 1.9% transgender. Depression, anxiety and stress scores derived from DASS-21 measures completed at baseline were used to characterise participant's level of common mental disorder symptomatology. Scores on the DASS-21 subscales revealed the mean level of depression (±SD) among participants was 10.01 ± 6.25, the mean level of anxiety was 6.47 ± 4.41, and the mean level of stress was 10.23 ± 5.05, indicating that for all three psychological states, the study sample had a moderate level of symptom severity.

Table 1. Baseline characteristics for participants

Variable	**Mean (SD)**	**Percentage**
Age	20.7 (3.2)	
Sex (female %)		77.0
Transgender (%)		1.90
Common Mental Disorder Measures		
DASS-21 Depression	10.01 (6.25)	
DASS-21 Anxiety	6.47 (4.41)	
DASS-21 Stress	10.23 (5.05)	

Note: N = 53 for Australian young adults between 18 to 25 years of age
Abbreviations: DASS-21 = Depression Anxiety Stress Scale - 21 items; SD = standard deviation.

Linked app sources

At the commencement of the study, 44% (*n* = 53) participants had downloaded and linked various activity monitoring apps and connected their app through the API software for eight months duration (see Table 2). Descriptive results showed that among the 53 participants, a total of 64 app sources had been linked, and 17% (*n* = 9) participants linked more than one app. The app source linked by the highest number of participants, and also with the most frequent number of observations was the *Moves* app with 31 participants and 1270 observations, indicating this was the most commonly popular choice of activity app among the group of participants. The app with the least amount of engagement by participants was the *Strava* app, with only one participant sharing 16 observations (see Table 1).

However, in relation to the amount of daily movement detected, participants spent the most time in movement with the wearable device *Fitbit* (M = 2.43, SD = 3.01), and the least amount of movement was detected by the smart phone app *Healthkit* (M = 0.00, SD = 0.00). In total, among the 53 participants, there were 3,276 daily observations of activity engagement from seven app sources combined, with an average of 1.49 (SD = 2.42) hours spent of activity duration. Collapsing the source into only two dichotonomous categories (i.e. either app based or device based), showed almost an equal split, with 54.44% (*n* = 1783) of the observation entries as app based sources (shown in Table 2). There were no significant correlations observed between DASS-21 scores and the number of activity engagement observation entries.

Table 2. Participant engagement, frequency and daily duration of activity tracker apps and devices

App Source	Number of participants	Total observation days	Daily duration (Mean, SD)
Wearables			
Fitbit	17	1515	2.43(3.01)
Garmin	2	221	2.41(2.57)
Apple Apps			
Healthkit	2	61	0.00(0.00)
Misfit	2	31	0.89(1.77)
Moves	31	1270	0.45(0.66)
Myfitnesspal	9	162	0.31(0.47)
Strava	1	16	0.66(0.08)

Note: N = 53 for Australian young adults between 18 to 25 years of age
Abbreviations: SD = standard deviation

4.2 Inferential statistics

Difference in daily duration between physical activity tracker sources

An independent samples *t*-test was conducted to compare the differences in daily durations of recorded physical activity between the two sources of data collection (wearable devices versus smartphone apps). Results revealed a positively, statistically significant difference in daily durations between the sources of data collection (*t*-test = 25.4, $p<0.001$). Specifically, it was found that detected daily durations (i.e. movement) was greater from wearable devices when compared with smartphone apps. No significant differences were observed between the different variants of DASS scores (i.e. depression, anxiety, stress) and source type.

Missing data between first and last observed day

To study temporal characteristics of missing data during the observation period, only those participants who had used an app or wearable device for more than three days were included. Consequently, ten participants were excluded as they had less than three daily observations of activity engagement available for analysis. As such, the final analysed cohort consisted of 43 participants. On average, these participants had 74.88 (SD = 80.23) observation days collected over the eight month intervention period. There was a positive, linear correlation between the number of days containing recorded observations with the number of days of missing data in between earliest and last day with recorded observations among the analysed cohort ($r = 0.67$, $p<0.001$). This finding indicates people who provided data for observation on a higher number of days over the period of 200 days, also had proportionately more days of missing data.

Of the 43 participants, 25.58% ($n = 11$) participants were missing more than 20% of data between the first and last available observation day (a period of 200 days). An independent samples *t*-test comparison with the group of participants ($n = 32$) who had the most amount of available observational data (i.e. missing less than 20%), showed no statistically significant relationship between the amount of missing data and DASS-21 depression, anxiety and stress scores.

Volatility in daily activity duration

Entropy techniques were used to compute volatility in daily activity duration over five different discretized time series for the 32 participants who had less than 20% observations missing was computed. Entropy, as defined by Shannon's theory (Robinson, 2008), characterises the degree of underlying randomness of a given variable. Variables with small entropies have a high level of predictability and hence a low level of randomness, whereby in contrast, variables with large entropies correspond to low levels of predictability and high levels of randomness (Vaseghi, 2008). A correlation analysis revealed none of the five entropy indexes demonstrated a significant relationship with any of the DASS-21 measures (see Table 3). However, when participants were grouped into two cohorts: 1. *less than 45 days measurement* *trajectories* (n = 16), and 2. *greater than 45 days* (n = 16), a positive, significant correlation between DASS-21 anxiety subscale scores and entropy of those with over 45 days measurements (r = 0.58, p = 0.02) was observed, indicating that entropy might be sensitive to the length of recorded observations.

Table 3: Correlations Between Entropy and DASS-21 Measurements.

	Depression		**Anxiety**		**Stress**	
Groups	***r***	***p***	***r***	***p***	***r***	***p***
All (n = 32)	-0.02	0.91	0.17	0.38	0.03	0.88
3-45 days (n = 16)	-0.43	0.11	-0.50	0.06	-0.35	0.21
45+ days (n = 16)	0.52	0.05	0.58	**0.02***	0.25	0.36

Note: Results are shown only for entropies calculated using random imputations as same correlations were observed for different imputation techniques.
Abbreviations: DASS-21 = The Depression Anxiety Stress Scale - 21 items; n = number of participants; p = statistical significance at p<.05 level; r = correlation coefficient

5. Discussion

At present, there is no standardised continuous monitoring system to track common mental health, symptom-related behavioural indicators, in real-time. This lack of progress, with the continuation of fixed-time monitoring, has contributed to an escalating incidence of common mental disorders worldwide (WHO, 2017). In this study, we examined commonly available activity tracker apps and wearables to continuously and passively collect and analyse behavioural indicators of clinically-validated symptoms of common mental disorders, specifically depression, anxiety and stress.

Among a sample of 53 young Australian adults with a moderate level of psychological distress, they were willing to use commercially available activity tracker apps for eight months and share their personalised data with researchers. Our findings concur with other recent studies in the area from the US (Grajales et al., 2014; Patankar, 2014), which show that a large percentage of individuals suffering with psychological distress are willing to share personal information from smartphones to help researchers learn more about mental illness. Furthermore, acceptance for researchers to track daily activity patterns using smartphone sensors has also been found for severe and serious mental illness populations (Ben-Zeev et al., 2015).

While further collaborative evidence is indeed needed, particularly from randomised controlled intervention trials, there is plausible potential for mobile sensing platforms to be used as a feasible and scalable alternative for scientific research in the area of mental health. Given long-term maintenance of smartphone app and wearable device usage has been shown to be associated with factors such as aesthetics, lifestyle compatibility, self-choice, ease of set-up, and a clear value to the user (Dennison et al., 2013), future research should continue to explore the efficacious difference between interventions providing dedicated smartphone apps to participants versus choosing their own commercially available apps.

Other findings from this study showed that a greater amount of daily activity duration was detected from wearable devices (e.g. *Fitbit* wristbands and watches) over smartphone apps. There are a few potential explanations that may clarify this finding, of which foremost, may owe to the sheer differences in the features and functions between wearables and smartphone apps. Indeed, a recent prospective study conducted by Naslund and colleagues (2016) investigated the acceptability of using *Fitbit Zip* wearable devices and smartphones apps to support a lifestyle intervention among a sample of participants with a serious mental illness and obesity for six months. Using mixed methods analysis, results from this study showed that all participants (100%) reported they found the *Fitbit Zip* wearable device useful, easy to use, were highly satisfied, and most importantly would continue to use it in the future. In contrast, when synching the *Fitbit Zip* wirelessly through Bluetooth directly to the companion *Fitbit* mobile application, participants reported several challenging issues that pertained to a much lower satisfaction level, and in some cases ceasing the use of it all together, including: poor phone reception which affected the performance of the mobile app, unfamiliarity with the technology, screen size being too small, short battery life, difficulty trying to exercise while carrying a smart phone, particularly when running.

Another potential explanation may relate to the level of accuracy achieved by wearable devices and smartphone apps for detection of physical activity with various intensities. In the small handful of studies which have investigated differences between exercise intensities (e.g. low, moderate, high/vigorous), results have been mixed depending on the device and type of physical activity. It has been found that accuracy captured in response to different levels of physical activity intensity may lead to erroneous results, as there may be both underestimation and overestimation of feedback (Dooley et al., 2017). In the present study we did not distinguish between the levels of activity intensities, as the primary focus of the study was more focused around user interaction and the feasibility of using sensing technology for obtaining mental health indicators. Hence, it is possible that the activity duration observations of participants in this study may have been over- or underestimated due to differences in sensitivity and specificity soundness between wearable devices and smartphone apps.

Finally, different types of algorithms are inherently employed in wearable devices versus smartphone apps. As such, even if activity duration is essentially the same, a person may register different values. This is also a reasonable explanation for the findings in this study. It will be important for future research to not only establish the feasibility and user acceptances of activity tracker sensing technology, but also to establish the psychometric validity associated with its use.

Other interesting finding found in this study was that while overall entropy indexes was not significantly related to any of the DASS-21 measures, in the sub-group of participants with over 45 days of observation data, we found an increase in randomness in daily activity durations (entropy) was significantly associated with worse anxiety symptoms. This can be interpreted as, the more anxiety distress participants were experiencing, the less stable functioning routines became over time. Or put another way, increasing volatility entropy is a function of high randomness and low predictability of customary physical activity in people with moderate levels of anxiety symptomatology. Therefore, variations in activity duration over time in individuals with psychological distress may be a potential context for intervention.

The results of this study should be interpreted in light of several limitations. Firstly, while the findings may be of interest to clinicians and patients as a potential means to track and alert to changes in mental symptoms in real time, given the observational nature of the study, the present findings do not assume a causal link between any of the behavioural physical activity indicators and clinical symptoms. Secondly, although we assessed depression, anxiety and stress using a well-validated measure, it was self-report. Hence, inherent bias may be accompanied with its use. We also utilised a convenience sample from the community as opposed to a clinically diagnosed sample. As such, the results may lack generalisability to the wider population. Third, while we used ecological momentary assessment to continuously monitor physical activity duration throughout the study, we did not distinguish the location of the movement (e.g. at home, the gym, work, shopping centre) or the level of activity intensity. Gathering this type of data from users would allow for an enhanced understanding of how people spend their time at various locations, such as whether they have not been able to leave the house in a number of days, are mainly sedentary, spending large amounts of time before or after typical work hours, or increasingly spending more and more time at the gym. All of which demonstrate potential lifestyle factors that may be associated with an increase in psychological distress, including depressive or anxious symptoms.

In summary, continuous monitoring using commercial apps and wearables, to identify, track and collect behavioural indicators of mental functioning, as a resource to help clinicians augment clinical care is viable (Harari et al., 2016). The use of such digitally, innovative approach could also empower patient health self-management. The findings from this study provide a basic "proof of concept" justification for the ability to use a continuous, digital monitoring approach with young adults experiencing a moderate level of psychological distress. As future research continues to evolve with this approach, the challenges in privacy, measurement, adherence, and clinical integration will need to be addressed before it can be broadly adopted within a clinical context, and self-management tool. In addition, we propose that research interventions which enable participants to choose their own smart phone apps and voice their needs, rather than focus on retrofitting smartphone apps and fitness wearables for medical applications, will encourage the development of devices more attuned to mental health needs. This will place greater emphasis on a user’s experience in tracking their own health, potentially improve overall user engagement, and increase ecological validity with inferred outcomes.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors contributions

A.K. formulated, wrote and drafted the manuscript and the final version, contributed to study design; N.B prepared and performed data collection, screening, analysis and interpretation of the data, formulation of the study design, provided feedback of the manuscript. All authors have approved the final version of the manuscript.

Conflict of interest

None.

References

Australian Department of Health, 2005. 1.4 - National Mental Health Plan 2003-2008. http://www.health.gov.au/internet/publications/publishing.nsf/Content/mental-pubs-n-infopri2-toc~mental-pubs-n-infopri2-pt1~mental-pubs-n-infopri2-pt1-4 (accessed 24 October 2017).

Alvarez-Lozano, J., Osmani, V., Mayora, O., Frost, M., Bardram, J., Faurholt-Jepsen, M., et al. 2014. Tell me your apps and I will tell you your mood: correlation of apps usage with bipolar disorder state. Paper presented at the Proceedings of the 7th International Conference on PErvasive Technologies Related to Assistive Environments, Rhodes, Greece.

Ben-Zeev, D., Scherer, E.A., Wang, R., Xie, H., Campbell, A.T., 2015. Next-Generation Psychiatric Assessment: Using Smartphone Sensors to Monitor Behavior and Mental Health. Psychiatr Rehabil J. 38, 218-226.

Beiwinkel, T., Kindermann, S., Maier, A., Kerl, C., Moock, J., Barbian, G., et al. 2016. Using Smartphones to Monitor Bipolar Disorder Symptoms: A Pilot Study. JMIR Ment Health, 3, e2.

Bidargaddi, N., Musiat, P., Makinen, V.P., Ermes, M., Schrader, G., Licinio, J., 2016. Digital footprints: facilitating large-scale environmental psychiatric research in naturalistic settings through data from everyday technologies. Mol Psychiatry. 22, 164.

Capela, N.A., Lemaire, E.D., Baddour, N., 2015. Feature selection for wearable smartphone-based human activity recognition with able bodied, elderly, and stroke patients. PLOS ONE. 10, e0124414.

Carlier, I.V., Meuldijk, D., Van Vliet, I.M., Van Fenema, E., Van der Wee, N.J., Zitman, F.G., 2012. Routine outcome monitoring and feedback on physical or mental health status: evidence and theory. J Eval Clin Pract. 18, 104-110.

Collins, P.Y., Price, L.N., Pringle, B.A., Developing and implementing collaborative research in global mental health: the NIMH Hubs. Lancet Glob Health. 2, S50.

De Mello, M.T., Lemos Vde, A., Antunes, H.K., Bittencourt, L., Santos-Silva, R., Tufik, S., 2013. Relationship between physical activity and depression and anxiety symptoms: a population study. J Affect Disord. 149, 241-246.

del Rosario, M.B., Redmond, S.J., Lovell, N.H., 2015. Tracking the Evolution of Smartphone Sensing for Monitoring Human Movement. Sensors (Basel). 15, 18901-18933.
Dennison, L., Morrison, L., Conway, G., Yardley, L., 2013. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. J Med Internet Res. 15, e86.

Dooley, E.E., Golaszewski, N.M., Bartholomew, J.B., 2017. Estimating Accuracy at Exercise Intensities: A Comparative Study of Self-Monitoring Heart Rate and Physical Activity Wearable Devices. JMIR Mhealth Uhealth. 5, e34.

Field, A., 2009. Discovering Statistics Using SPSS: SAGE Publications.
Goldberg, D., 1994. A bio-social model for common mental disorders. Acta Psychiatr Scand Suppl. 385, 66-70.

Grajales, F.D., Clifford, P., Loupos, S., Okun, S., Quattrone, M., Wicks, S.P., Henderson, D., 2014. Social networking sites and the continuously learning health system: A survey. http://nam.edu/wp-content/uploads/2015/06//SharingHealthData.pdf (accessed 25 October 2017).

Grunerbl, A., Muaremi, A., Osmani, V., Bahle, G., Ohler, S., Troster, G., et al. 2015. Smartphone-based recognition of states and state changes in bipolar disorder patients. IEEE J Biomed Health Inform, 19, 140-148.

Harari, G.M., Lane, N.D., Wang, R., Crosier, B.S., Campbell, A.T., Gosling, S.D., 2016. Using Smartphones to Collect Behavioral Data in Psychological Science: Opportunities, Practical Considerations, and Challenges. Perspect. Psychol. Sci. 11, 838-854.

Lovibond, S.H., Lovibond, P.F., 1995. Manual for the Depression Anxiety Stress Scales, second ed. Psychology Foundation, Sydney.

Mayora, O., Arnrich, B., Bardram, J., Dräger, C., Finke, A., Frost, M., et al. 2013. Personal health systems for bipolar disorder Anecdotes, challenges and lessons learnt from MONARCA project. Paper presented at the 2013 7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops.

Mojtabai, R., Olfson, M., Mechanic, D., 2002. Perceived need and help-seeking in adults with mood, anxiety, or substance use disorders. Arch Gen Psychiatry. 59, 77-84.

Naslund, J.A., Aschbrenner, K.A., Marsch, L.A., Bartels, S.J., 2016. The future of mental health care: peer-to-peer support and social media. Epidemiol Psychiatr Sci. 25, 113-122.

New Zealand Guidelines Group, 2008. Identification of Common Mental Disorders and Management of Depression in Primary Care. An Evidence-based Best Practice Guideline. New Zealand Guidelines Group, Wellington.

Osmani, V., 2015. Smartphones in Mental Health: Detecting Depressive and Manic Episodes. IEEE Pervasive Comput, 14, 10-13.

Parkitny, L., McCauley, J., 2010. The Depression Anxiety Stress Scale (DASS). J Physiother. 56, 204.

Patankar, T., 2014. Accenture interactive: four in five consumers cite privacy concerns for wearable tech adoption, according to 2014 State of the Internet of Things Study. http://newsroom.accenture.com/news/four-in-five-consumers-cite-privacy-concerns-for-wearable-tech-adoption-according-to-2014-state-of-the-internet-of-things-study-from-accenture-interactive.htm. (Accessed 23 October 2017).

Pratt, L.A., Brody, D.J., 2014. Depression in the U.S. household population, 2009-2012. NCHS Data Brief. 172, 1-8.

World Health Organization, 2017. Depression and Other Common Mental Disorders: Global Health Estimates. Geneva. http://apps.who.int/iris/bitstream/10665/254610/1/WHO-MSD-MER-2017.2-eng.pdf (accessed 30 October 2017).

World Health Organization, 2016. Investing in treatment for depression and anxiety leads to fourfold return. http://www.who.int/mediacentre/news/releases/2016/depression-anxiety-treatment/en/ (accessed 24 October 2017).

Vaseghi, S.V., 2008. Advanced Digital Signal Processing and Noise Reduction: Wiley.

Vogl, G., Ratnaike, D., Ivancic, L., Rowley, A., Chandy, V., 2016. One Click Away? Insights into Mental Health Digital Self-help by Young Australians. Sydney: EY and ReachOut Australia.

Zhang, S., McClean, S., Scotney, B., Chaurasia, P., Nugent, C., 2010. Using duration to learn activities of daily living in a smart home environment. Paper presented at the 2010 4th International Conference on Pervas.