Depression Assessment Based on Technology using Smartphone Data

Rui Miguel Rodrigues Barbosa
Extended Abstract of thesis to obtain the Master of Science Degree in Biomedical Engineering
Instituto Superior Técnico, Lisboa, Portugal

Abstract

The general definition of psychiatric disorders is based on subjective symptoms, reported by the patient, which normally forms diagnostic patterns. Those symptoms are generally difficult to detect, characterize or quantify mostly because mental health doesn’t have biomarkers to analyse. Smartphones offer the promise of collecting behavioural data unobtrusively, in situ, as it unfolds in the course of daily life. Data can be collected from the onboard sensors and other phone logs embedded in today’s off the shelf smartphone devices. These data permit fine grained, continuous collection of people’s social interactions, daily activities and mobility patterns, which means that we can collect valuable information about psychological states. It was collected objective behavioural data (acceleration, ambient light level, battery level, GPS coordinates and phone logs) with Ethica Health app. Not only objective, but also subjective data (PHQ-9 scores, stress barometer and sleep diary) were collected from all participants with that app to represent the traditional methods of depression diagnosis. Having the two types of data, it was possible to correlate them, and the results were very optimistic, where people more depressed showed more stress ($r=0.87$), less activity ($r=0.89$) and less location variability ($r=0.85$). A desktop application to treat and visualize data as well as provide useful metrics from the data in depression diagnosis was also designed and fulfilled its purpose. This proof of concept study showed that smartphones can be used as instruments for unobtrusive collection of behavioural data that are associated with depression, something that can be revolutionary in this area.

Keywords: Mental Health, Depression, Diagnosis, Objective Data, Smartphone, Passive sensing

1. Introduction

1.1 Mental Illness – Depression

1.1.1 Epidemiology

Between 2005 and 2015, the number of people living with depression increased in 18.4% [1] which is due to the growth of the global population allied with a proportionate increase in the age groups that depression is more prevalent that happened in this period. In Portugal, the numbers are even more worrying. According to official numbers [2]:

- More than one fifth of Portuguese people suffers from some mental disorder;
- At least 8% of the population has depression;
- It is estimated that there are 400 thousand Portuguese people between the ages of 18 and 65 who suffer from depression each year;
- In 70% of suicide cases the cause is depression and every 8 hours a Portuguese die of suicide.

Having all these numbers in mind, it is evident that depression is one of the biggest illnesses of our time and for that reason it is very important that its
prevention and treatment is done as efficiently as possible.

1.1.2 Definition
Depressive perturbations are characterized by sadness, lack of interest or pleasure, guilty and low self-esteem feelings, change of sleeping and eating patterns, tiredness and low concentration levels. It can be lasting or recurrent and can lead to a severe negative impact on people's capacity of doing their daily activities [1]. In the last instance, depression can lead to suicide.

There are multiple factors and risks that can explain depression development. In fact, the responsible is not only one factor, instead there are several factors that interact with each other (genetic, psychological, environmental and biological factors) to bring on depression (figure 1). So, for that reason, every person that suffers from depression is different and that is the main difficulty when physicians are dealing with this disorder.

![Figure 1 - Common causes for depression. Source: Psychological Health Center of Excellence.](image)

1.1.3 Diagnosis
As mentioned before, depression is a very complex disorder and for that reason it is still difficult to treat. Current common practice for treatment selection is an educated guess approach, in which physicians prescribe their approved therapies in a stepwise manner [3]. There is already a lot of evidence based clinical guidelines to manage depression that can be already a support to specific treatment recommendations, but the symptom heterogeneity present in the diagnosis of depression creates a significant barrier to it. And now we are beginning to get into the main problem that exists nowadays associated with mental disorders, the way that the diagnosis is made and the tools that exist to make it in the best way possible.

In the consultation, the physician conducts a complete diagnostic evaluation based on depression family history or other mental illness, evaluate the symptoms and duration of them that can last for weeks, months and even years and the episodes may occur only once in a lifetime or may be recurrent or chronic. In severe cases, it could be last forever.

As we saw, what the physicians wants to know is always related to how the patients behave in their daily live. Mostly they want to know about patient's sociability, physical activity and sleeping patterns. So, now the problem is how it is possible to measure patient's behaviour in their daily lives.

1.2 Technology based diagnostic support tools in psychiatry

1.2.1 Wearable Devices
In the past three decades, computational power has increased a lot, the storage space prices have dropped and electronic components are miniaturised in a way that a few years ago it was impossible to even imagine that [4]. With that, researchers started to think about a device that can monitor people in their daily lives, which is the key behind psychiatry diagnosis. The only possible way to monitor behaviour is asking people to wear a device that collects useful data to this purpose and here enters the big importance of wearable devices in psychiatry. Researchers started to think in many possible wearable devices equipped with multiple
sensors and this has been an area of great interest from the researcher point of view \cite{4}.

Studies with wearable sensors are divided in two sections, the first one is related to those who make the collection in a laboratory environment, where people have to complete a serious of tasks and then researchers do their analysis based on the data they have collected from the behavioural response to those tasks. The bias in this first section is obvious, people aren't acting normally as they do in their normal lives. It is important to collect objective behaviour as people play out in the context of their natural lives. The second one is related to those who do their study with people using the wearable device in their daily lives, but they also have issues. These methods have been difficult and time consuming to use and very intrusive for the participants, and if the person is not feeling as if he has nothing clinging to his body, he is not behaving naturally, which is a very important bias to consider. The best alternative to being as unobtrusive as possible would be to deal with a wearable device that people already wears almost every time. A great candidate to eliminate all those biases referred before is the smartphone, a wearable device that is very powerful and most of the people who use it are unaware of it.

1.2.2. Smartphone as a Behavioural Collection tool

According to Marktest Telecommunications Barometer \cite{5} in 2018 there was nearly 6.8 million smartphone owners in Portugal, representing close to three quarters of all mobile phone owners. Over the last few years, the possession of this type of mobile phone has been on an upward trend, rising from 32.5% in 2012 to 73.9% in December 2017\cite{5}. Since this market is already so huge and the trend is to increase in the next years, working with this wearable device seems to be perfect. Adding to that, these phones are very sensor rich, increasingly computationally powerful and, as we saw, their ubiquity can be very helpful to us since it provides unparalleled access to people daily lives in an unobtrusive, dense and continuously way. Another factor that contributes to this is the capacity that smartphones have to query people about their subjective psychological states with notifications, which is perfect because with that we have not only objective data but also subjective data that is already validated by years of medical clinic. So, using the smartphone as a diagnostic support tool seems to be an idea with lots of potential.

Smartphones already come equipped with all the sensors that are needed to obtain the best behaviour description possible. The information that they give are divided into three categories: Social interactions, Daily activities and Mobility patterns Working with smartphone sensor’s data it is possible to describe social interactions as Monsivais D. et al \cite{6} did. In that study, the authors explored mobile phone calling activity and concluded that the onset and termination of the resting pattern of urban humans follow sun progression and they also showed that calling activity period follows the same dynamics as solar midnight. Monitoring call and SMS messages logs for frequency, duration and unique persons contacted in incoming and outgoing interactions it is possible to know easily people’s social interactions. Adding to that microphone and Bluetooth sensors it is also possible to have an approach of face to face interactions. Daily activities can also be described with data from smartphone sensors, as physical activity (using the accelerometer) or sleeping patterns (with exterior light, accelerometer and phone usage logs). Mobility patterns are also very important to monitor people’s routines, which is very important
to get the best possible behaviour description. These patterns included people’s duration of time spent in various places (home or work), the distance travelled in each time period and of course their routines. All of this can be captured with GPS, accelerometer, exterior light or Wi-Fi sensors. With all those embedded sensors, smartphone now offers an easy data acquisition offering a lot of opportunities to continuously monitor patients in their natural environments due to its ubiquity.

Our proposal is to get a proof of concept of smartphone sensors utility in depressed people and to build an app that will help the physician with objective measurements about his patients behaviours, never forgetting that our goal is to help the physician to obtain the best diagnosis possible and never to replace his important role. Without a physician sensitivity the efficiency of a mental health diagnosis will greatly decrease, so our goal is to add one more diagnostic aid tool to the panoply that the physician already has to, along with all traditional methods, increase his diagnostic accuracy.

2. Methods
2.1 Study Design

The total number of participants was 10, on which 4 were male, 5 were female and 1 didn’t want to say. 5 participants were aged between 21-30 years, 3 between 31-40, 1 between 41-50 and 1 between 51-60. 6 had a degree and the other 4 had compulsory education. Regarding the employment situation, 2 of them were students and the other 8 were employees. Each participant had the application ‘Ethica’ from Ethica Health \(^8\) installed on his smartphone and was associated with our study in the app, a process that was accompanied by the creators of the study. After all the installation process of the required software, the participant was asked to answer a demographic questionnaire and then answered other three questionnaires that had two different time intervals, two of them daily and the other weekly. They completed two short questionnaires a day (one in the morning and one before bed) called sleep diaries and filled out the stress barometer once a day at night for the entire 31-day study period. At the beginning of the study, each participant answered the PHQ-9 \(^7\) questionnaire and then they repeat it weekly until the last day of the study. After completing the first questionnaires, the data extraction from the smartphone sensors started. This data collection was continuous throughout the study. Data from accelerometer, GPS, ambient light intensity, battery level and duration and frequency of calls and SMS was extracted (no information was recorded about the content of SMS or calls). Concluding, apart from questionnaires responses, no further procedure was required except to behave as usual in their normal life. Participants were instructed to enable the GPS sensor at all time during the study. Due to privacy reasons, Ethica Health also provides an option to turn off data collection for a period of one hour at a time and if the sensors are accidentally or not turn off, the app automatically notifies the participants that collection is paused.

2.2 Diagnostic Support Tool

After collecting the data, it is very important to have some program that can handle data processing and visualization and that returns some useful metrics in the depression diagnosis context from the data and a desktop application that handles all of that was built. The home page of it is represented in figure 2, where it has buttons to each of the sensor data that was used in this study (accelerometer, GPS, Exterior Light, Phone Logs and battery level) and the path of each one.
After downloading the data, it is important to preprocess the data. The most problematic data is related to the accelerometer because adding to the noise situation that is common for other sensors, this data combines acceleration due to sensor motion but also due to gravity. To solve this problem, it was used a low pass and high pass filter. The goal is to filter out the portion of accelerometer data caused by gravity from the portion of the data that is caused only by the sensor motion which is the important one for us. The effect of low pass and high pass filtering on each time series is to create gravitational and body acceleration data in different directions.

In figure 3 there is represented plots of the original accelerometer magnitude signal, also of the magnitude of the low-pass filter values which are the ones related to gravity effects and the bottom one is the plot of acceleration magnitude related to the action of both low and high pass filters, i.e. the values that are related only with sensor motion.

Then, the page that appears in our desktop app is the one represented in figure 4. This page has two buttons that redirects to two different things. The ‘Plots’ button takes you to a data visualization tool and the ‘Metrics’ button shows a page of some meaningful metrics in depression context that were developed.

If we click on the ‘Plots’ button, the page that appears is the one represented at figure 5. Visualizing the data is a very important part of this app because it gives to the physician the possibility to see the sensor’s data in a timeseries. This can be very helpful for them because in that way they can make their own analyses and get their conclusions analysing it for themselves. For instances if they want to see the mobility patterns of a specific day or week or if the patient makes many social interactions or not at a specific time period this is possible with this tool. This possibility can be a great help in diagnosing depression.
the data from that day from the sensors combination that he wants, including all of them together and the trajectory that the person did in that day.

Instead of clicking on button ‘Plots’, if you click on ‘Metrics’ button, the page that appears is the one on figure 6. Here there are three different buttons: ‘Social Interactions’, ‘Mobility Patterns’ and ‘Daily Activity’.

![Figure 6 – Page of metrics information.](image)

Of course, letting the physician get their own analyses by data visualization is great, but they also want some meaningful metrics that automatically show up simply by clicking in a button. It is more immediate and requires less time to get meaning from the data. Regarding the possible metrics, an important one is people’s sociability. As we saw before, knowing if a person is sociable or not is essential in depression diagnosis. Being sociable or not is essential in depression as we saw before. For that reason, the button ‘Social Interactions’ shows: the number and duration/length of incoming and outgoing calls and SMS, number of calls missed, number of messages failed and the number of unique calls. This last one is very important because having 20 calls from 1 person has a very different meaning of having 20 calls from 20 persons. The same thing works for SMS obviously. Therefore, with this button it is possible to see the sociability of people during the study time.

Regarding the mobility patterns, this button shows how people mobilized by calculating the travelled distance and location variability with GPS data. With longitude and latitude coordinates, it is possible to know person’s travelled distance by using the Haversine formula. This formula allows us to calculate the distance between two points in a sphere, as it is the case with the Earth. This formula was implemented to obtain the daily distance travelled by people based on GPS data.

The last button is the one related to daily activity. To show daily activity, we used accelerometer data because it is a motion sensor very common in all smartphones and the subject doesn’t have the option to turn off the sensor, which is the case with GPS for example. The accelerometer data was summarized into a high activity variable by calculating the percentage of time at which the summed variance of the device’s acceleration on the three axes was above a defined threshold, that was 10 m/s². So, when the summed variance exceeded that value, it counted as high activity sample. As the other two metrics, these percentages were aggregated to the day level in order to provide an approximate measure of daily activity.

### 3. Results

None of the participants had issues dealing with the Ethica Health application on their smartphones. In fact, all of them report that the interaction with it was simple and didn’t disturb their normal use with the smartphone. The surveys that worked as a baseline were completed by all of them and that is a proof of the simplicity of dealing with the app. An important factor that distinguished the participants were the PHQ-9 scores. All of them completed that questionnaire weekly till the end of the study. Regarding that, 20% of them showed no depressive symptoms, 20% showed mild depression, 50% reported moderate depression, 10% moderately severe depression and none of
them reported severe depression. This separation between the participants were fundamental to all the results that will be show next.

Beginning with the daily stress scores, it was model the relationship between daily stress and PHQ-9 scores. The results showed that people with higher PHQ-9 scores had also higher values of daily stress values \((r=0.87)\). This relationship was established between the mean of the all daily stress scores in a week and the PHQ-9 scores that were obtained weekly. The participants with more stress are also the ones that showed more depression severity. It is also important to refer that in the weeks that the participants showed less score in PHQ-9 questionnaire was also in the weeks that reported less daily stress levels.

The application that was built worked fine with all the data and the plots that were used in the analysis were all of them obtained with the buttons from figure 5.

![Figure 7 - Chart of all the sensors. Acceleration (x, y and z m/s2), Phone logs events, battery level and intensity of exterior light (lux) over time.](image)

In figure 7, it is represented the chart of all the sensors over time. This is the most important one, because in this one it is possible to combine all the sensors information to get the most accurate approximation of person behaviour. Looking at figure 7, we can see that at night the ambient light level and the accelerometer values are nearly zero and the phone was on charge. So, the combination of these three sensors values makes a great method to predict sleeping patters. Having that in mind, it was compared the sleep duration of some participants between two methods, the first with these sensors and the second one with the reports that were filled out by the participants in the sleep diaries. The results that were obtained were good, because having the reports as a baseline, the sleeping patterns calculated with the sensors information had only an error by approximately 1 hour, which is actually not bad due to the few participants and few time that this study had. The difference between night and day were clearly notice as we can see in figure 7, during the day people are more active, more exposed to light, have more social interactions and waste more battery due to increased phone usage.

Regarding the metrics, it was also correlated with PHQ-9 scores. All the metrics were obtained with the app that was built by clicking on buttons that are presented at figure 6. Starting with mobility patterns, an example is represented at table 1.

<table>
<thead>
<tr>
<th>Date</th>
<th>Distance_Travelled</th>
<th>Sum_Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>2019-08-29</td>
<td>0.5148</td>
<td>0.0000</td>
</tr>
<tr>
<td>2019-08-30</td>
<td>37.5890</td>
<td>0.0011</td>
</tr>
<tr>
<td>2019-08-31</td>
<td>135.1566</td>
<td>0.0011</td>
</tr>
<tr>
<td>2019-09-01</td>
<td>159.2831</td>
<td>0.0014</td>
</tr>
<tr>
<td>2019-09-02</td>
<td>123.1133</td>
<td>0.0008</td>
</tr>
<tr>
<td>2019-09-03</td>
<td>134.7305</td>
<td>0.0005</td>
</tr>
<tr>
<td>2019-09-04</td>
<td>46.2797</td>
<td>0.0002</td>
</tr>
</tbody>
</table>

The distance travelled and location variability correlated very well with PHQ-9 scores, having people with less PHQ-9 scores (people with less tendency to be depressive) higher values in distance travelled and higher location variability \((r=0.85)\). The difference between workday and weekend was also notice in this metric since the values for weekend presented higher values in location variability due to routine behaviours in workdays.
Regarding the phone logs information, this metric report all the social interactions that a subject does during all the study time. Table 2 represents an example of the output of the button ‘Social Interactions’ that is showed in figure 6.

Table 2 - Phone Logs information.

The correlation between social interactions and PHQ-9 scores didn’t exist. For that reason, it is not possible to conclude that social isolation is necessarily a depressive signal, it clearly depends on the person. Despite that, people with higher values of PHQ-9 showed fewer social interactions variability, having less unique contacts that other subjects with less scores of PHQ-9 questionnaire. But it is very important to know people background and old habits to be possible to make a robust conclusion with this objective data. As we saw, information can have two distinct meanings, because being antisocial may represent a tendency to depression, but the opposite is also true.

The last metric that was used was daily activity and an example of it is represented in table 3. This metric is also very important in depression diagnosis and a correlation between these values and PHQ-9 scores was obtained. Participants that showed higher values in weekly activity (calculated as a median of all the daily activities of the week) were the ones that reported less scores in PHQ-9 questionnaire \((r=0.89)\). Since activity is an important feature to correctly characterize depression symptoms this is an important conclusion.

So, with this app it was possible characterize social interactions, mobility patterns and activity with objective data and compared it with subjective methods (traditional methods) both collected via smartphone, getting significant correlations between them.

4. Discussion and Future Work

This proof of concept study showed that smartphones can be used as instruments for unobtrusive collection of behavioural data that are associated with depression. None of the participants had issues dealing with the Ethica Health app, which was very good for the reliability of the results because the goal was to act as more natural as possible. The results were very optimistic since in almost all features that were created the correlation was good.

The app that was designed to treat and visualize data as well as provide useful metrics from the data in depression diagnosis fulfilled its purpose. All results associated with smartphone sensor data came from that application, both charts and metrics. With a tool like this, a physician can make an analysis of people’s behaviour with higher efficiency because this tool works with objective data (which is an innovation in depression diagnosis) where in the visualization part he can see the sensor combination that he wants, with the
possibility of choosing a specific day and zoom in or out at any time of that day. From these graphs and metrics, it was possible to verify the correlation between objective and subjective data.

With data from accelerometer, ambient light sensor and battery level it was possible to calculate people’s sleeping patterns, which was then compared to the ones reported by the participants in their sleep diaries. The results were goods, because having the reports as a baseline, the sleeping patterns calculated from the sensors information had only an error by approximately 1 hour. To decrease this error, the ideal for calculate sleeping patterns using only objective data is by combining 6 types of features: light, phone usage, motion and acoustic records. This combination is perfect because it monitors all the things that normally happen when we go to sleep, as darkness, phone recharge, phone static state and silence.

Regarding mobility patterns that were calculated with GPS coordinates data and daily activity computed with accelerometer data, the correlations between them and PHQ-9 scores were very significant. People with less location variability and less weekly activity tend to be more depressive. Even in the same participant, the weeks that were more ‘depressive’ were the ones that had less mobility and activity, which clearly indicates that these sensors are very important in monitoring people’s behaviour with smartphone. The difference between workday and day off is also very important because in workdays people tend to follow a routine due to their work responsibilities, so when we are analyse these data it is essential to have that in consideration.

With plots of Phone Logs, it was possible to see if people were very sociable in a specific day and at what time of day, they were more interactive with their phone, something that can be very helpful for physician depression diagnosis. With that it was possible to see participant’s social interactions, but a correlation between it and PHQ-9 scores wasn’t find. This happened because social interactions are clearly a personalized issue, some people like more to interact than others. For that reason, when we are analysing this metric, is very important to know previously people background to see if the social interactions pattern changed or not, which demands a longer study period. This is an example of the importance of personalized medicine, which is clearly the future of medicine since it will increase its effectiveness.

This study has several limitations. Despite all the relevant findings, the number of participants was very small, but it was representative of the broader population in terms of their demographic characteristics, education, level of functioning, or willingness to engage in smartphone monitoring. All of them worked with the Ethica Health app with no problem. The study duration was also small because for some people one month isn’t enough to accurately detect mood changes, but the participants were not monolithic in terms of their behaviours and mental health status. Some were very active while others were not, some reported no mental health difficulties while others endorsed more severe symptoms of depression, which was great for our conclusions. One of the objectives of this study was to have participants with real diagnosed depression and it was unfortunate that the Beatriz Ângelo hospital ethics committee accepted the consent form too late, otherwise patients with depression were present in this study. It is a limitation because PHQ-9 is only a screening tool that tells you if you have an above-average chance of having depression and is not enough to diagnose depression. Future research should use
that people with actual depression in order to get more significant conclusions.

As already mentioned, an app like the one created in this study can be an extraordinary tool for mental health physicians. Not only in consultation environment, where they can make their own analyses looking at the data and the metrics while asking the patient questions, but also outside the consultation. When a person goes to a psychiatric appointment, the doctor makes his diagnosis and there will be no further contact between them until the next appointment. In fact, currently in Portugal this is an important issue regarding the waiting time for psychiatric consultation. Most of the time, people wait weeks, months or even a year for a psychiatric appointment [9]. This tool can be useful for the waiting time problem as the doctor can continue to monitor the patient remotely. With this technique, the doctor can verify whether the treatment he has done is having the desired effect or not without the need for consultation, something that would make the need for so many consultations greatly decreased, and the follow-up of each person would be better, and more personalized.

In future the ideal is to build an application that receives all the sensors information and within the same application it calculates sleeping patterns with all the combination data sensors that is required as well as mobility patterns, daily activity and social interactions in a personalized way. That means that the app previously knows how each person normally behaves and ideally if a risk behaviour has been taking, the app automatically will send an alert to the physician and notifies the patient that the behaviour he is taken is risky. Data must always be protected, and people must have control over their own data. To incorporate this into normal psychological consultations at health centres, health organizations need to be interested in this new tool, which is believed to be very soon due to the growing interest in this subject that is nowadays a hot topic when we are talking about fusion between technology and medicine.

As technology evolves, it is exciting to imagine a future in which individuals who could benefit from additional support have sensor-enabled data collection systems installed on their smartphones and calibrated to their individual needs. The trend towards personalized medicine is strengthening and discreet monitoring techniques such as this fit perfectly. In the future, the diagnosis of mental health in general and specifically depression will be completely different for sure. All new technology-based approaches will be common to physicians as a tool just like any other tool they use today, and this will make the diagnosis of depression more accurate.

5. Acknowledgement

This document was written and made publically available as an institutional academic requirement and as a part of the evaluation of the MSc thesis in Biomedical Engineering of the author at Instituto Superior Técnico. The work described herein was performed at the Institute for Systems and Robotics of Instituto Superior Técnico (Lisbon, Portugal), during the period March-September 2019, under the supervision of Prof. João Miguel Raposo Sanches. The thesis was co-supervised at Hospital Beatriz Ângelo by Dr. Miguel Constante.

6. References


