Implementation of automatic speech analysis for early detection of psychiatric symptoms: What do patients want?


Abstract

Psychiatry is in dire need of a method to aid early detection of symptoms. Recent developments in automatic speech analysis prove promising in this regard, and open avenues for implementation of speech-based applications to detect psychiatric symptoms. The current survey was conducted to assess positions with regard to speech recordings among a group (n = 675) of individuals who experience psychiatric symptoms. Overall, respondents are open to the idea of speech recordings in light of their mental welfare. Importantly, concerns with regard to privacy were raised. Given that speech recordings are privacy sensitive, this requires special attention upon implementation of automatic speech analysis techniques. Furthermore, respondents indicated a preference for speech recordings in the presence of a clinician, as opposed to a recording made at home without the clinician present. In developing a speech marker for psychiatry, close collaboration with the intended users is essential to arrive at a truly valid and implementable method.

Keywords:
Automatic speech analysis
Patient involvement
Early detection
App-based monitoring

1. Introduction

An estimated 10.7% of the world population suffered from a psychiatric disorder in 2017 (Ritchie and Roser, 2018), imposing a high burden on mental health care and society. Part of this can be attributed to the pressing lack of a reliable method for early detection of symptoms. Detecting signals of emerging psychiatric disorders or relapse would enable timely intervention and prevention of advanced stages of a disorder. However, at present patients often already have full-blown symptoms by the time they visit a mental health care professional.

Previous studies suggest that the information encapsulated in spontaneous speech might serve as a marker for symptom severity in psychiatric disorders, facilitating differential diagnosis (Lott et al., 2002) and the prediction of emerging symptoms in individuals who are considered at-risk (for recent reviews, see Low et al., 2020; Corcoran and Cecchi, 2020). Furthermore, aberrant speech has been related to structural brain measures as underlying psychiatric symptomatology (De Boer et al., 2020a,b). Importantly, speech disturbances can be externally and objectively assessed, as opposed to some other psychiatric phenomena (e.g., delusions) (Tan and Rossell, 2020). This opens up new avenues for tracking a person’s mental health, by way of monitoring changes in their speech over time. Cutting-edge developments in machine learning technology using speech samples can offer a valuable way to improve such a marker for psychiatry (De Boer et al., 2020a,b; Robin et al., 2020), and form ideal candidates for implementation in, for example, websites or smartphone-based apps. Advantages of tracking spontaneous speech (as opposed to, for example, self-reported symptom scales) are its ecological validity and relative immunity to influences of response bias. Moreover, app-based mental health tracking has the added advantage of high temporal precision, low effort required by the patient, and the option of providing real-time results and feedback (Colombo et al., 2019). First attempts to track mental health via speech recordings show promising results (Braun et al., 2016; Arevian et al., 2020), which will have to be replicated and extended by future studies.

While the potential of a speech marker in psychiatry is clear, an important unanswered question is whether patients with psychiatric problems are in fact receptive to the idea of mental health tracking based on recorded speech. Before implementation of a speech application for...
psychiatry, it is important to learn about the wishes and concerns of the intended users, and to explore facets that would improve acceptance and user-friendliness. To this end, we conducted a survey in collaboration with MIND – a Dutch collaborative platform for organizations of people with mental health problems and their caregivers. With this survey, we assessed the respondents’ position with regard to smartphone-based speech recording, website-based recording, and speech recording in the presence of a mental health care professional.

2. Methods

The development of the survey questions was a collaborative effort with input from clinicians (SB, JB, IS), voice and app-technicians (AV, JW), and a patient expert (PG) (Table 1). For questions 1–4, respondents were given the chance to further explain their answers, but this was not mandatory. If respondents had indicated a preference for a recording during a conversation with a clinician (question 5) they received a follow-up question asking for their motivation, and likewise in case they had indicated a preference for answering neutral questions (question 6).

The survey was distributed online amongst members of the MIND mental health panel who indicated to experience psychiatric problems. The data were collected between the 8th of October and the 3rd of November 2020. In this period, 675 members of the MIND panel completed the online survey after giving informed consent, of whom the majority (80%) was female, with a mean age of 49 (range 28–64 years). Of the respondents, 63% indicated to have been diagnosed with a depressive disorder; 41% with an anxiety disorder; 14% with bipolar disorder; and 9% with a psychotic disorder.

3. Results

The majority of the respondents (66%) indicated to welcome the idea of having their speech recorded to help their mental welfare (Fig. 1.1). Another 24% showed reservations, but indicated to be interested if concerns regarding privacy would be taken away. More than half of the respondents (54%) indicated they would consider downloading an app to manually record their speech in light of their mental health (Fig. 1.2). Doubts with regard to privacy were given as primary concern in those (31%) who reported no interest in an app to manually record their speech. Some respondents suggested that it could be a burden to manually record their speech, especially in light of their mental health issues. However, 49% of the respondents indicated that because of privacy concerns they would not download a speech app on their smartphone if it would record their speech automatically (Fig. 1.3). Part of the respondents indicated they might (29%) or surely would (21%) make use of a website to record their speech to track their mental health (Fig. 1.4). Another 30% said they would rather not use such a website, again mostly because of privacy concerns. A number of respondents commented that they did not feel digitally apt enough to deal with smartphone apps or websites.

Interestingly, the majority of the respondents (68%) favored a speech recording during a conversation with their clinician over a recording that would be made without the presence of a health care professional (Fig. 1.5). Motivations for this choice included saving time and effort by having a recording run during a conversation that the respondent was going to have anyway, as well as the appeal of a speech recording in the presence of somebody they feel safe with. When asked whether neutral questions or mood-related questions would be favored during a recording, the majority of respondents (56%) indicated to have no preference (Fig. 1.6).

4. Discussion

In sum, the intended users –individuals who experience psychiatric problems– are open to the option of speech-based mental health monitoring. Most participants prefer recording of a conversation with their clinicians over making recordings for this purpose by themselves. Furthermore, respondents of the survey highlighted their concerns with regard to privacy. This chimes with recent reports on the topic of ethical considerations associated with automatic analyses of data. Howe and Elenberg (2020) highlight the challenges accompanying the quick advancements in computational analyses of large datasets. Specifically, they identify ethical risks which call for stronger regulations. An example is that participants should be given the opportunity to get a complete understanding of how their data are used. Hendrikoff et al. (2019) report that psychiatric patients express concerns with regards to data security in mobile health tracking, although generally they believed it could be an enrichment for medical practice. This shows that in developing and implementing speech-recording applications for psychiatry, it is of paramount importance to take measures to ensure the safety of the data and to clearly inform intended users about the use of their data and privacy issues.

Some limitations of this study should be noted. The fact that the majority of the sample was female detracts from its generalizability. Furthermore, psychiatric diagnoses were self-reported and had not been externally confirmed, nor did we acquire measures of symptom severity. In the current study, comorbidity and unbalanced group sizes prevented us from differentiating results between diagnoses, but previous studies suggest acceptability and cooperativeness may relate to diagnosis and symptom severity. Ben-Zeev et al. (2015) showed that within a group of patients with a schizophrenia spectrum disorder, outpatients had no concerns with regards to privacy issues in collecting smartphone-based data, while one-third of inpatients did. A study comparing a group of patients predominantly diagnosed with a psychotic disorder with a group predominantly diagnosed with an anxiety or mood disorder, found the former to be less likely to download a mental health app and to feel less comfortable with all of its features (Torous et al., 2018). This further highlights the importance of understanding patients’ attitudes towards speech-based mental health monitoring in light of their clinical diagnosis, and could be the focus of future studies.

As a final note, the current covid-19 crisis forms an alarming underscore of the necessity to detect and prevent psychiatric symptoms, especially in the face of limited possibilities to physically visit a health care institution. Tracking mental health at home by using speech analysis that is based on the latest machine learning developments may prove an excellent candidate to fulfill this need. By gaining an impression of patients’ position regarding speech recordings, researchers, clinicians, and patients work together in reaching their common goal: diminishing the occurrence and associated burden of psychiatric disorders.

Author statement


