

Original Paper

# Psychiatric Profiles of eHealth Users Evaluated Using Data Mining Techniques: Cohort Study

Jorge Lopez-Castroman<sup>1,2,3,4\*</sup>, MD, PhD; Diana Abad-Tortosa<sup>2\*</sup>, PhD; Aurora Cobo Aguilera<sup>5,6</sup>, MS; Philippe Courtet<sup>1,4,7</sup>, MD, PhD; Maria Luisa Barrigón<sup>8,9</sup>, MD, PhD; Antonio Artés<sup>3,5,6</sup>, PhD; Enrique Baca-García<sup>2,3,8,9,10,11,12,13</sup>, MD, PhD

<sup>1</sup>Institute of Functional Genomics, CNRS-INSERM, Montpellier, France

<sup>2</sup>Department of Psychiatry, Nimes University Hospital, Nimes, France

<sup>3</sup>CIBERSAM, Madrid, Spain

<sup>4</sup>University of Montpellier, Montpellier, France

<sup>5</sup>Department of Signal Theory, Universidad Carlos III de Madrid, Madrid, Spain

<sup>6</sup>Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, Spain

<sup>7</sup>Department of Psychiatric Emergency and Acute Care, Lapeyronie Hospital, University of Montpellier, Montpellier, France

<sup>8</sup>Universidad Autonoma de Madrid, Madrid, Spain

<sup>9</sup>Department of Psychiatry, Hospital Universitario Fundación Jiménez Díaz, Madrid, Spain

<sup>10</sup>Department of Psychiatry, University Hospital Villalba, Villalba, Madrid, Spain

<sup>11</sup>Department of Psychiatry, University Hospital Infanta Elena, Valdemoro, Madrid, Spain

<sup>12</sup>Department of Psychiatry, University Hospital Rey Juan Carlos, Mostoles, Madrid, Spain

<sup>13</sup>Universidad Católica del Maule, Talca, Chile

\*these authors contributed equally

**Corresponding Author:**

Jorge Lopez-Castroman, MD, PhD

Institute of Functional Genomics, CNRS-INSERM

39 Ave Charles Flauhault

Montpellier, 34093

France

Phone: 33 4996145

Email: [jorgecastroman@gmail.com](mailto:jorgecastroman@gmail.com)

## Abstract

**Background:** New technologies are changing access to medical records and the relationship between physicians and patients. Professionals can now use e-mental health tools to provide prompt and personalized responses to patients with mental illness. However, there is a lack of knowledge about the digital phenotypes of patients who use e-mental health apps.

**Objective:** This study aimed to reveal the profiles of users of a mental health app through machine learning techniques.

**Methods:** We applied a nonparametric model, the Sparse Poisson Factorization Model, to discover latent features in the response patterns of 2254 psychiatric outpatients to a short self-assessment on general health. The assessment was completed through a mental health app after the first login.

**Results:** The results showed the following four different profiles of patients: (1) all patients had feelings of worthlessness, aggressiveness, and suicidal ideas; (2) one in four reported low energy and difficulties to cope with problems; (3) less than a quarter described depressive symptoms with extremely high scores in suicidal thoughts and aggressiveness; and (4) a small number, possibly with the most severe conditions, reported a combination of all these features.

**Conclusions:** User profiles did not overlap with clinician-made diagnoses. Since each profile seems to be associated with a different level of severity, the profiles could be useful for the prediction of behavioral risks among users of e-mental health apps.

(*JMIR Ment Health* 2021;8(1):e17116) doi: [10.2196/17116](https://doi.org/10.2196/17116)

**KEYWORDS**

mental disorders; suicide prevention; suicidal ideation; data mining; digital phenotyping

## Introduction

The development of new technologies shows promise for causing a revolution in the way chronic diseases are followed and treated [1]. In the past few decades, we have seen the following two major technological changes directly connected to the availability of medical information: (1) introduction of electronic health records in most health care facilities, and (2) accessibility to portable devices capable of acquiring information about their users. Both systems are already being used to enhance communication between health providers and final users and to improve the overall performance of health care. Indeed, public and private entities are massively investing in the development of web-based platforms or smartphone apps through which patients can organize their medical agenda, have access to all or part of their medical records, provide their input, and join their medical referents [2].

It seems reasonable to believe that the follow-up of persons with mental illness will be improved if eHealth systems lead to an increased interaction with health care providers. Our group and others have shown that electronic assessment is feasible with proper adaptation [3]. Efficient monitoring may prompt health responses in cases of emergency [4], inform accurately about real-life behaviors between medical appointments, reduce unnecessary visits, and sustain therapeutic decisions [5]. Other parameters, such as biomarkers and input from close relatives, can be added to the monitoring system. This kind of ecosystem already exists [6], and a growing body of evidence has shown that eHealth tools improve treatment outcomes in terms of engagement, symptom improvement, well-being, and self-care [7-13]. Their combination with machine learning techniques has also shown positive impacts on the diagnosis, prediction, and prevention of several diseases, such as cancer [14-16].

Despite these advances, the majority of potential users, such as elderly persons with low educational levels [17], seem to be unenthusiastic about e-mental health tools [18]. Utilization rates have been associated with the characteristics of health professionals [19], but there is little knowledge about the kinds of patients who use e-mental health, how they become users, and what are their patterns of use. Young age, high education, and dissatisfaction with the health care system might be common features among eHealth users [20]. We do not know if this profile also applies to patients with mental disorders, but their digital phenotype [21] is likely to contain valuable information for clinicians and providers alike [22,23]. Alterations in the patterns of use could help clinicians to detect pathological or risky behaviors and individual needs, and increase treatment efficiency.

In this article, a nonparametric latent feature model based on the Indian Buffet Process (IBP) explores the response patterns of 2254 psychiatric outpatients to a web-based questionnaire. We aimed to define specific profiles according to response patterns and then link profiles with psychiatric disorders. The study will thus specifically describe patients with psychiatric diagnoses who have used an eHealth application at least once. We hypothesize that data mining techniques, such as IBP, can be used to associate information from different questionnaires

and assessments in a plausible model that could serve ultimately to plan health care delivery.

## Methods

### Sample

Participants were recruited from psychiatric outpatient facilities in the catchment area of Fundacion Jimenez Diaz, a University Hospital in Madrid, Spain. This hospital is part of the National Health Service and provides medical coverage to about 850,000 people. From May 2014 onwards, all clinicians working at the six mental health centers of the catchment area received specific training and were encouraged to use the MEmind Wellness Tracker systematically in their clinical activity. A total of 2254 patients signed up on the MEmind platform and completed the assessment, and they were subsequently included in the study. The assessment comprised the collection of information about sociodemographic features and diagnoses. Participants also filled up a short questionnaire. For this study, we used broad inclusion criteria. Every patient attending psychiatric consultations independent of diagnosis was considered. Thus, all clinicians in the catchment area were instructed to propose the use of the web application to every outpatient they saw with no restriction whatsoever regarding their diagnoses or their clinical statuses. The total number of outpatients who consulted during the study period was 30,808.

For the purpose of this study, we included only participants who voluntarily accessed the application and responded to an open-text field. We made this choice to select proactive participants who completed most of the questions at the user end. We noted a missing data rate of 12%, which resulted from the sum of clinical missing data and a lack of completeness of the questionnaires at the user end of the application.

This study was performed in agreement with the ethic requirements of the Declaration of Helsinki (World Medical Association, 2013) and was approved by the Institutional Review Board of the University Hospital Fundación Jiménez Díaz (Madrid, Spain). All participants provided written informed consent to participate in the study.

### Assessment

#### Questionnaires

The data set consists of 23 questions from the following three different questionnaires: (1) a brief day assessment related to sleep quality, appetite, medication intake, aggressiveness, and suicidal behavior (six items); (2) the Who-5 Well-Being Index [24] (five items); and (3) the ninth version of the General Health Questionnaire [25] (12 items). All these questionnaires are short self-reported measures of current mental well-being. All items are yes-or-no questions, followed by the degree of agreement reported on a Likert scale (0 to 100 points). Although participants could repeat the assessment, only data from the first report were included in the model.

#### Clinical Diagnoses

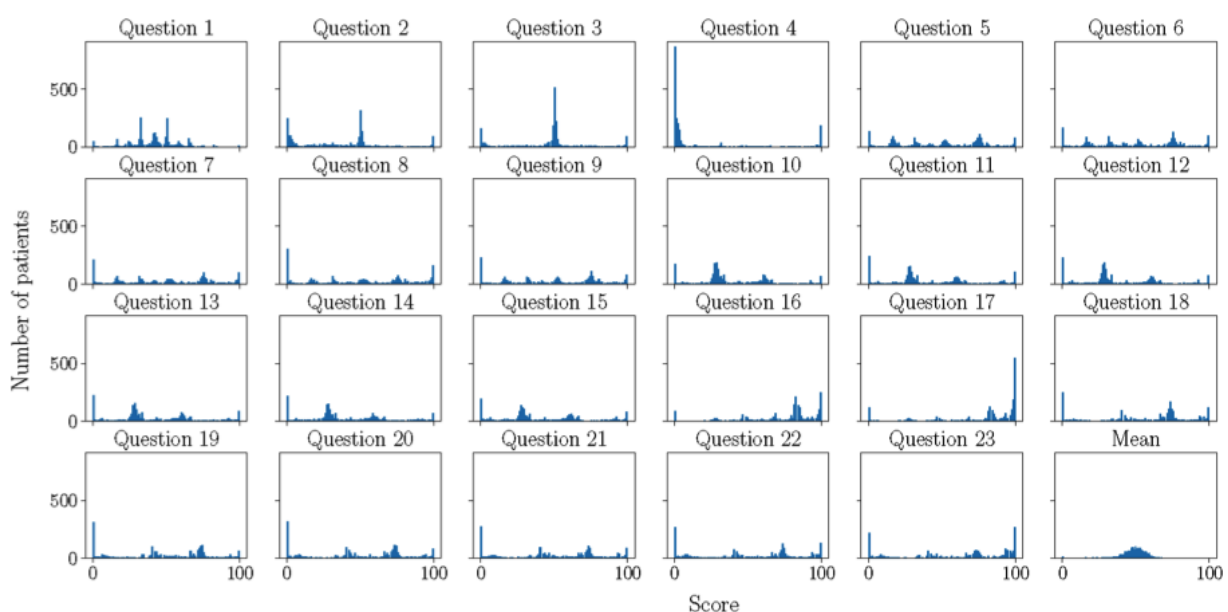
Diagnostic coding was based on the International Statistical Classification of Diseases and Related Health Problems (ICD-10) [26]. Thus, diagnoses of mental disorders were

classified into 10 groups (F0 to F9) according to ICD-10 (Multimedia Appendix 1). The corresponding physician coded the diagnosis for each patient and completed the clinical global impression (CGI) scale [27], which reflects the global functioning of a patient according to the view of the clinician on a scale (0 to 7 points). The CGI scale provides a summary measure accounting for patient history, psychosocial factors, behavior, and the impact of symptoms on the patient's ability to function (Multimedia Appendix 2).

### Data Processing

First, the scores for the items with a positive valence in the questionnaire data set (items 1 to 15) were inverted. In this way,

**Figure 1.** Histograms of scores provided by eHealth users to each of the 23 questions of the mental well-being questionnaire. Scores range from 0 to 100. The last histogram presents the average score for all questions.



The clinical records of the participants provided a second source of data. These records included sex, age, clinical diagnoses, and CGI values. CGI values presented missing data, so subsequent analyses including this variable were carried out with a total sample of 2000 participants. For analyses involving clinical diagnoses, the total sample was 1787 participants. All patients with missing data were excluded from this part of the data modeling. Comorbid diagnoses were also examined when present.

### Data Modeling

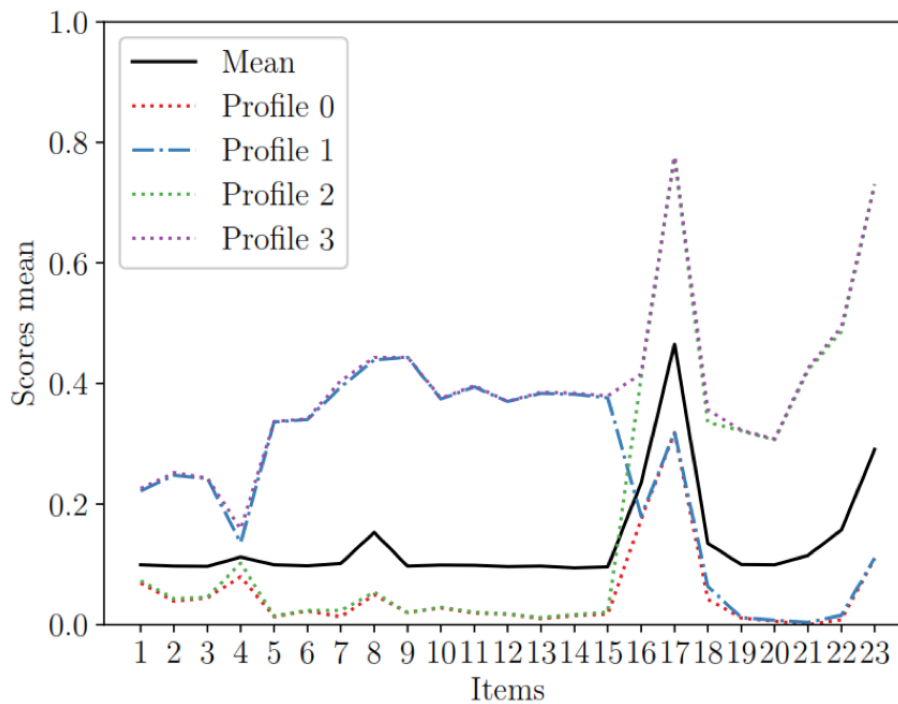
We applied the *Sparse Poisson Factorization Model* (SPFM) to model the data. The SPFM is based on the IBP [30], a nonparametric probabilistic method that proposes a sparse analysis of the variables. The input data for the model must be binary or categorical. The SPFM decomposes the input matrix into the following two nonnegative and disperse matrixes: matrix  $Z$  and matrix  $\beta$ . Latent factor sets can be calculated from them. The binary  $Z$  matrix represents the number of active factor sets for each patient (Multimedia Appendix 4). The  $\beta$  matrix weights

a higher score for any item of the questionnaire indicated poorer mental health. Second, we dichotomized every item score using a specific threshold in order to code the top 10% scores with the value “1” and the remaining 90% with the value “0” (Multimedia Appendix 3). The use of a centesimal scale increases the sensibility of the questionnaire. Responders tend to avoid extreme values unless they identify completely with them [28], but extreme responders do not seem to be affected by the length of the response scale [29]. By using the highest scores, we made sure that only the extreme responders were separated. The histograms of dichotomized scores for the items are shown in Figure 1.

the contribution of each factor set to each item of the questionnaire. Each factor set is characterized by precise values on the 23 questions. A higher weight ( $\beta$ ) of a factor set for an item is associated with a greater probability to find a high score in that item when that particular factor set is active. The SPFM also estimates a *bias term*, a factor set that is present in all the patients of the sample [31]. The bias term is the “default” situation of an eHealth user and represents a profile shared by all patients that is independent of any additional feature. In that sense, the bias term allows the machine learning function to be shifted to better fit the data in a similar way as done by the y-intercept in an equation.

Different profiles have been obtained by applying the basic clustering method K-means [32]. The procedure classifies a given data set through a certain number of clusters fixed a priori. This method, applied on the  $Z$  matrix, associates data with similar characteristics into different clusters by using centroids. Thus, it allows clustering patients who show similar activation of their factor sets (Figure 2).

**Figure 2.** Item scores according to the four patient profiles and their averages for the whole sample.



## Results

### Sample Description

The sample involved 2254 patients, including 1184 (52.53%) women, 795 (35.27%) men, and 275 (12.20%) patients with missing data on sex. The mean age was 52.0 years (SD 15.1). Medical reports about the patients showed a CGI mean score of 2.95 (SD 1.95) (Multimedia Appendix 5) with a high percentage of participants scoring 3 (mildly ill; 844/2000, 42.20%) or 4 (moderately ill; 632/2000, 31.60%) and only a few scoring 6 or 7 (severely ill or extremely ill; 10/2000, 0.50%). According to the ICD-10 criteria, participants with mood disorders (F3; 347/1787, 19.43%), stress-related disorders (F4; 962/1787, 53.82%), and adult personality disorders (F6; 178/1787, 9.96%) represented most of the sample (Multimedia Appendix 6).

### Data Classification

The SPFM latent model analysis found the following three components in the assessment: one *bias term* and two *factor*

*sets*. Both the bias term and factor sets involved groups of items of the questionnaire that are particularly informative. The bias term is present for all patients and reflects a common behavioral pattern. On the other hand, factor sets 1 and 2 are based on subsets of answers with high informational value to discriminate patients. Factor sets 1 and 2 can be present or absent for a particular patient. They are present if the corresponding subset of responses has a high score (value “1”: highest 10% scores) and are absent if the corresponding subset of responses has a low score (value “0”: remaining 90% of the scores).

The K-means algorithm applied on the Z matrix established four different patient profiles according to the presence of none, one, or both factor sets. Profile 0 presents only the bias term, profile 1 presents the bias term plus factor set 1, profile 2 presents the bias term plus factor set 2, and profile 3 presents the bias term and each factor set. The number of patients in each profile is shown in Table 1. In Table 2 and Multimedia Appendix 7, we can appreciate how each profile responds differently to the questionnaire.

**Table 1.** Number of patients in each profile.

Profile	Number of patients	Factor set
0	1113	Bias term (0)
1	480	0+1
2	616	0+2
3	45	0+1+2

**Table 2.** Average score for each item in the self-reported questionnaire of current mental well-being according to the  $\beta$  matrix of factor sets.

Questions	Profile 0	Profile 1	Profile 2	Profile 3
	Bias term (0)	Factor set 0+1	Factor set 0+2	Factor set 0+1+2
1. How many hours did you sleep today? (from 0 to 12) <sup>a</sup>	0.0689	0.1525	0.0041	0.1567
2. Quality of sleep <sup>a</sup>	0.0390	0.2089	0.0040	0.2129
3. Do you have appetite? <sup>a</sup>	0.0446	0.1980	0.0009	0.1990
4. Do you take your medication? <sup>a</sup>	0.0795	0.0566	0.0225	0.0791
5. I felt joyful and with good mood <sup>a</sup>	0.0133	0.3228	0.0006	0.3235
6. I felt peaceful and relaxed <sup>a</sup>	0.0223	0.3181	0.0012	0.3193
7. I felt active and robust <sup>a</sup>	0.0130	0.3811	0.0103	0.3915
8. I felt awake, fresh, and rested <sup>a</sup>	0.0496	0.3894	0.0041	0.3936
9. My daily life has many interesting things <sup>a</sup>	0.0199	0.4236	0.0000	0.4236
10. Have you been able to keep focus on the tasks you did? <sup>a</sup>	0.0276	0.3467	0.0013	0.3481
11. Have you felt that you have a useful role in life? <sup>a</sup>	0.0187	0.3756	0.0018	0.3774
12. Have you felt able to make decisions? <sup>a</sup>	0.0173	0.3529	0.0002	0.3531
13. Have you enjoyed regular activities from daily life? <sup>a</sup>	0.0100	0.3739	0.0016	0.3756
14. Have you felt able to cope with your issues? <sup>a</sup>	0.0149	0.3674	0.0016	0.3691
15. Do you feel reasonably happy taking into account the circumstances? <sup>a</sup>	0.0174	0.3591	0.0030	0.3622
16. Do you feel aggressiveness?	0.1745	0.0053	0.2362	0.2415
17. Do you have suicidal thoughts?	0.3177	0.0012	0.4573	0.4585
18. Have you had worries interfering with your sleep?	0.0422	0.0210	0.2931	0.3142
19. Have you felt constantly overwhelmed or tense?	0.0109	0.0006	0.3108	0.3115
20. Have you felt unable to overcome your troubles?	0.0060	0.0015	0.3005	0.3021
21. Have you felt unhappy or depressed?	0.0001	0.0037	0.4213	0.4251
22. Have you lost self-confidence?	0.0083	0.0071	0.4784	0.4855
23. Have you felt worthlessness?	0.1097	0.0010	0.6198	0.6208

<sup>a</sup>The scores from these items were inversed during data processing.

## Patient Profiles

The bias term was associated with high scores in suicide thoughts and aggressiveness (items 5 and 6), as well as feelings of worthlessness (item 23). All patients in our sample shared the features of the bias term, but about half of them (n=1141) also presented one or two different factor sets. Those presenting factor set 1 were included in profile 1, which was characterized by the absence of positive mood, low sleep quality, low energy, and feelings of loss of control (items 1-3 and 5-15). Patients presenting factor set 2 were included in profile 2, which was characterized by intense suicidal thoughts, aggressiveness, intense feelings of depression and worthlessness, low self-confidence, and worries interfering with sleep (items 16-23). All these characteristics were simultaneously active in patients with profile 3, who presented simultaneously with both factors. No statistical differences were found between the profiles

regarding the distribution of age or sex ( $F_3=1.391$ ,  $P=.24$  and  $\chi^2_3=0.56$ ,  $P=.90$ ).

## Association Between Patient Profile and Medical Evaluation

After modeling the data, we compared CGI scores and clinical diagnoses between profiles. The results showed that the CGI scores were higher than the mean in profile 1 (3.2, SD 1.27), with the largest percentage of participants evaluated with a score of 4 (192/453, 42.4%). For profile 2, the CGI scores were lower than the mean (2.78, SD 1.28), with a high percentage of participants evaluated with a score of 3 (256/557, 45.9%). Results in profile 3 were not compared given the low number of patients.

Most diagnoses fell within the F4, F3, and F6 categories in each profile and in the total sample, corresponding with affective disorders, neurotic and stress-related disorders, and disorders

of adult personality and behavior. The distribution of participants with profile 1 was similar for all the types of diagnoses. However, profile 2 seemed to be more frequent among patients

with diagnoses of schizophrenia and psychological, behavioral, and emotional disorders with onset in childhood/adolescence (F2: 43/90, 48%; F8: 3/5, 60%; and F9: 14/38, 37%; [Table 3](#)).

**Table 3.** Distribution of patient profiles according to the main ICD-10 diagnostic categories for psychiatric disorders (F0-F9).

Category	Profile, n (%)			
	0	1	2	3
F0	7 (50.0)	3 (21.4)	2 (14.3)	2 (14.3)
F1	26 (45.6)	13 (22.8)	18 (31.6)	0 (0.0)
F2	37 (41.1)	10 (11.1)	43 (47.8)	0 (0.0)
F3	215 (49.6)	104 (24.0)	104 (24.0)	10 (2.3)
F4	614 (51.2)	269 (22.4)	290 (24.2)	26 (2.2)
F5	51 (48.1)	25 (23.6)	29 (27.4)	1 (0.9)
F6	109 (49.1)	58 (26.1)	51 (23.0)	4 (1.8)
F7	6 (66.6)	1 (11.1)	2 (22.2)	0 (0.0)
F8	2 (40.0)	0 (0.0)	3 (60.0)	0 (0.0)
F9	20 (52.6)	3 (7.9)	14 (36.8)	1 (2.6)

## Discussion

The data modeling approach we applied was able to discriminate four different profiles of patients based on the answers to a brief electronic questionnaire. All profiles shared a component associated with feelings of aggressiveness, worthlessness, and suicidal thoughts (bias term or profile 0), which seemed to be common among patients who used e-mental health tools [33], such as the MEmind Wellness Tracker.

Sex and age distributions showed very little variability across the profiles, facilitating comparisons between them. In addition to profile 0 (bias term, default pattern), three profiles were found based on the scores of different sets of questions. It is important to bear in mind that a factor set was classified as active only when the scores were in the top 10% of the corresponding items. For example, even if the item of low sleep quality is absent from profile 2, a patient with that profile could still have high scores in that item compared with the general population and thus have relatively low sleep quality.

Patients in profile 1 reported a lack of positive mood, low quality of sleep, low energy, feelings of loss of control, and difficulties to face problems. These symptoms could be reactive to life difficulties and partly due to a lack of coping skills. Patients in profile 2 presented high scores in depressive feelings, worries interfering with their sleep, feelings of being overwhelmed and unable to overcome troubles, low self-confidence, and feelings of worthlessness. This pattern seems to be related with a greater inward focus and depressive-like symptomatology. Interestingly, patients in profile 2 also reported the highest scores for suicidal thoughts and feelings of aggressiveness. Indeed, patients in profile 2 reported five of the 10 ICD-10 diagnostic criteria for a depressive episode, including disturbed sleep, depressive feelings, reduced self-confidence, ideas of worthlessness, and ideas of suicide [26]. Surprisingly, those in profile 2 were evaluated by their physicians as having a higher level of functionality (CGI) than those in profile 1, despite higher levels

of suicidal thoughts, aggressiveness, and depression in profile 2. This points to discordances between the medical assessment and the self-reported momentary assessment. Finally, profile 3 involved a small group of patients with high scores in all the items of the questionnaire. They shared the features of profile 1 and profile 2, and reported the most severely affected psychological state in our sample (the highest levels of distress).

Our study suggests that the analysis of data from electronic self-assessments can discriminate profiles or clusters of patients sharing similar clinical characteristics. These features do not seem to overlap with usual clinical diagnoses, since no differences were found in the prevalence of previous psychiatric diagnoses between profiles. Most patients in each profile received diagnoses in F4 (anxiety disorders) and F3 (mood disorders) ICD-10 categories, which were numerically the most common diagnoses in the sample. However, diagnoses of disorders with an onset during childhood and adolescence (eg, F8 and F9) and schizophrenia (F2) were overrepresented in profile 2. Profile 3 was particularly overrepresented among the small group of patients with organic mental disorders (F0) in the sample, which could implicate a more complex disease course. Interestingly, in a previous paper, we found that the assessments made by clinicians did not correlate well with patients' self-reports within 24 h of a clinical evaluation [34].

The presence of sporadic suicide thoughts can be relatively frequent in psychiatric patients, but eHealth apps could help to identify profiles with higher suicide risk, such as profile 2. Previous literature has suggested improvements in mood, well-being, anxiety, and self-awareness, as well as a higher adherence to treatments among users of eHealth apps [5,35-37]. Electronic assessment tools, such as the one used in our study, may support physicians to discriminate patients with high suicide risk in order to adjust their interventions.

Among the limitations of our study, we note the use of only baseline assessments and incomplete clinical information. The

described profiles might not be reflective of eHealth users who continue to use the app regularly. Besides, our intention was not to map the participants onto Diagnostic and Statistical Manual of Mental Disorders (DSM) or ICD categories but rather to identify symptomatic profiles that are not necessarily reflected in psychiatric diagnoses. This study was designed to explore the utility of a new method to classify e-mental health users, and it needs to be completed with follow-up data. Nonetheless, once the SPFM is trained, it will be possible to analyze changes in patient profiles during continuous assessment with several time points. It will also be possible to link the electronic assessment with medical records. Our results could help to select the most performing questions according to mental disorders or patient profiles, which, in turn, could be used to create shorter and more efficient questionnaires. We can see in our study that the question about medication intake had very low informative value.

There are still many concerns regarding e-mental health that need to be addressed. One of the main concerns reported by both professionals and users is related to the privacy, ownership, and responsible use of medical information [38,39]. This is one

of the major challenges that eHealth needs to address by means of privacy-preserving technologies [40]. Accessibility and difficulties to find reliable sources of medical information are also important concerns in the population, especially among older adults [41]. Medical professionals also have doubts about the capacity of online information to improve the knowledge of patients and have reported concerns regarding the capacity of telemedicine to enhance physician-patient bond [39]. All these concerns must be addressed in order to improve the acceptability and use of eHealth tools. A recent study suggested that there is still a low preference for the use of eHealth tools among the adult general population [17]. However, those who have already used eHealth apps usually feel confident to continue using them. Some studies have reported a sense of security and the existence of a relational bond between eHealth apps and patients with psychiatric diseases [42,43]. Our analyses show that machine learning can help to classify e-mental health users and provide clues for their diagnoses and, importantly, their needs in terms of treatment. If machine learning helps physicians to take clinical treatment decisions based on data, the social perception about available eHealth tools will certainly improve.

---

## Acknowledgments

This study received grant support from Instituto de Salud Carlos III (ISCIII PI13/02200; PI16/01852), Delegación del Gobierno para el Plan Nacional de Drogas (20151073), the American Foundation for Suicide Prevention (LSRG-1-005-16), the Ministerio de Ciencia, Innovación y Universidades (RTI2018-099655-B-I00; TEC2017-92552-EXP), and the Comunidad de Madrid (Y2018/TCS-4705, PRACTICO-CM).

---

## Conflicts of Interest

EBG and AAR designed the MEmind application. The authors declare no other conflicts of interest.

---

## Multimedia Appendix 1

List of categories from the ICD-10 Classification of Mental and Behavioral Disorders.

[\[DOCX File , 13 KB-Multimedia Appendix 1\]](#)

---

## Multimedia Appendix 2

Clinical global impression (CGI) scores: severity of illness.

[\[DOCX File , 12 KB-Multimedia Appendix 2\]](#)

---

## Multimedia Appendix 3

Dichotomized processing of data from the questionnaire scores.

[\[DOCX File , 14 KB-Multimedia Appendix 3\]](#)

---

## Multimedia Appendix 4

A binary Z matrix, presenting the number of active factor sets for each patient. Each line corresponds to a single patient. All patients present the bias term or factor set 0.

[\[DOCX File , 48 KB-Multimedia Appendix 4\]](#)

---

## Multimedia Appendix 5

Number of patients by clinical global impression (CGI) scores.

[\[DOCX File , 28 KB-Multimedia Appendix 5\]](#)

---

## Multimedia Appendix 6

Percentage of patients according to ICD-10 diagnoses.

[\[DOCX File , 26 KB-Multimedia Appendix 6\]](#)

## Multimedia Appendix 7

Average scores for each item of the self-reported questionnaire of current mental well-being according to the gamma distribution. The model is based on a  $\beta$  matrix. A higher weight ( $\beta$ ) of a factor set for an item is associated with a greater probability to find a high score in that item when that particular factor set is active.

[\[PNG File , 30 KB-Multimedia Appendix 7\]](#)

## References

1. Oldenburg B, Taylor CB, O'Neil A, Cocker F, Cameron LD. Using new technologies to improve the prevention and management of chronic conditions in populations. *Annu Rev Public Health* 2015 Mar 18;36:483-505. [doi: [10.1146/annurev-publhealth-031914-122848](https://doi.org/10.1146/annurev-publhealth-031914-122848)] [Medline: [25581147](https://pubmed.ncbi.nlm.nih.gov/25581147/)]
2. A digital revolution in health care is speeding up. *The Economist*. 2017. URL: <https://www.economist.com/business/2017/03/02/a-digital-revolution-in-health-care-is-speeding-up> [accessed 2020-12-29]
3. Barrigón ML, Rico-Romano AM, Ruiz-Gomez M, Delgado-Gomez D, Barahona I, Aroca F, MEmind Study Group. Comparative study of pencil-and-paper and electronic formats of GHQ-12, WHO-5 and PHQ-9 questionnaires. *Rev Psiquiatr Salud Ment* 2017;10(3):160-167. [doi: [10.1016/j.rpsm.2016.12.002](https://doi.org/10.1016/j.rpsm.2016.12.002)] [Medline: [28237432](https://pubmed.ncbi.nlm.nih.gov/28237432/)]
4. Yellowlees P, Burke MM, Marks SL, Hilty DM, Shore JH. Emergency telepsychiatry. *J Telemed Telecare* 2008 Sep;14(6):277-281. [doi: [10.1258/jtt.2008.080419](https://doi.org/10.1258/jtt.2008.080419)] [Medline: [18776070](https://pubmed.ncbi.nlm.nih.gov/18776070/)]
5. Bucci S, Schwannauer M, Berry N. The digital revolution and its impact on mental health care. *Psychol Psychother* 2019 Jun;92(2):277-297. [doi: [10.1111/papt.12222](https://doi.org/10.1111/papt.12222)] [Medline: [30924316](https://pubmed.ncbi.nlm.nih.gov/30924316/)]
6. Berrouguet S, Barrigón ML, Castroman JL, Courtet P, Artés-Rodríguez A, Baca-García E. Combining mobile-health (mHealth) and artificial intelligence (AI) methods to avoid suicide attempts: the Smartcrises study protocol. *BMC Psychiatry* 2019 Sep 07;19(1):277 [FREE Full text] [doi: [10.1186/s12888-019-2260-y](https://doi.org/10.1186/s12888-019-2260-y)] [Medline: [31493783](https://pubmed.ncbi.nlm.nih.gov/31493783/)]
7. Payne HE, Lister C, West JH, Bernhardt JM. Behavioral functionality of mobile apps in health interventions: a systematic review of the literature. *JMIR Mhealth Uhealth* 2015 Feb 26;3(1):e20 [FREE Full text] [doi: [10.2196/mhealth.3335](https://doi.org/10.2196/mhealth.3335)] [Medline: [25803705](https://pubmed.ncbi.nlm.nih.gov/25803705/)]
8. Thakkar J, Kurup R, Laba T, Santo K, Thiagalingam A, Rodgers A, et al. Mobile Telephone Text Messaging for Medication Adherence in Chronic Disease: A Meta-analysis. *JAMA Intern Med* 2016 Mar 1;176(3):340-349. [doi: [10.1001/jamainternmed.2015.7667](https://doi.org/10.1001/jamainternmed.2015.7667)] [Medline: [26831740](https://pubmed.ncbi.nlm.nih.gov/26831740/)]
9. Badawy SM, Barrera L, Sinno MG, Kaviany S, O'Dwyer LC, Kuhns LM. Text Messaging and Mobile Phone Apps as Interventions to Improve Adherence in Adolescents With Chronic Health Conditions: A Systematic Review. *JMIR Mhealth Uhealth* 2017 May 15;5(5):e66 [FREE Full text] [doi: [10.2196/mhealth.7798](https://doi.org/10.2196/mhealth.7798)] [Medline: [28506955](https://pubmed.ncbi.nlm.nih.gov/28506955/)]
10. Al-Durra M, Torio M, Cafazzo JA. The use of behavior change theory in Internet-based asthma self-management interventions: a systematic review. *J Med Internet Res* 2015 Apr 02;17(4):e89 [FREE Full text] [doi: [10.2196/jmir.4110](https://doi.org/10.2196/jmir.4110)] [Medline: [25835564](https://pubmed.ncbi.nlm.nih.gov/25835564/)]
11. Saslow LR, Mason AE, Kim S, Goldman V, Ploutz-Snyder R, Bayandorian H, et al. An Online Intervention Comparing a Very Low-Carbohydrate Ketogenic Diet and Lifestyle Recommendations Versus a Plate Method Diet in Overweight Individuals With Type 2 Diabetes: A Randomized Controlled Trial. *J Med Internet Res* 2017 Feb 13;19(2):e36 [FREE Full text] [doi: [10.2196/jmir.5806](https://doi.org/10.2196/jmir.5806)] [Medline: [28193599](https://pubmed.ncbi.nlm.nih.gov/28193599/)]
12. Naparstek J, Wing RR, Xu X, Leahey TM. Internet-delivered obesity treatment improves symptoms of and risk for depression. *Obesity (Silver Spring)* 2017 Apr;25(4):671-675 [FREE Full text] [doi: [10.1002/oby.21773](https://doi.org/10.1002/oby.21773)] [Medline: [28224713](https://pubmed.ncbi.nlm.nih.gov/28224713/)]
13. Haas K, Martin A, Park KT. Text Message Intervention (TEACH) Improves Quality of Life and Patient Activation in Celiac Disease: A Randomized Clinical Trial. *J Pediatr* 2017 Jun;185:62-67.e2 [FREE Full text] [doi: [10.1016/j.jpeds.2017.02.062](https://doi.org/10.1016/j.jpeds.2017.02.062)] [Medline: [28343658](https://pubmed.ncbi.nlm.nih.gov/28343658/)]
14. Obermeyer Z, Emanuel EJ. Predicting the Future - Big Data, Machine Learning, and Clinical Medicine. *N Engl J Med* 2016 Sep 29;375(13):1216-1219 [FREE Full text] [doi: [10.1056/NEJMp1606181](https://doi.org/10.1056/NEJMp1606181)] [Medline: [27682033](https://pubmed.ncbi.nlm.nih.gov/27682033/)]
15. Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. Machine learning applications in cancer prognosis and prediction. *Comput Struct Biotechnol J* 2015;13:8-17 [FREE Full text] [doi: [10.1016/j.csbj.2014.11.005](https://doi.org/10.1016/j.csbj.2014.11.005)] [Medline: [25750696](https://pubmed.ncbi.nlm.nih.gov/25750696/)]
16. Dreiseitl S, Ohno-Machado L, Kittler H, Vinterbo S, Billhardt H, Binder M. A comparison of machine learning methods for the diagnosis of pigmented skin lesions. *J Biomed Inform* 2001 Feb;34(1):28-36 [FREE Full text] [doi: [10.1006/jbin.2001.1004](https://doi.org/10.1006/jbin.2001.1004)] [Medline: [11376540](https://pubmed.ncbi.nlm.nih.gov/11376540/)]
17. Paslakis G, Fischer-Jacobs J, Pape L, Schiffer M, Gertges R, Tegtbur U, et al. Assessment of Use and Preferences Regarding Internet-Based Health Care Delivery: Cross-Sectional Questionnaire Study. *J Med Internet Res* 2019 May 16;21(5):e12416 [FREE Full text] [doi: [10.2196/12416](https://doi.org/10.2196/12416)] [Medline: [31099338](https://pubmed.ncbi.nlm.nih.gov/31099338/)]



18. Musiat P, Goldstone P, Tarrrier N. Understanding the acceptability of e-mental health--attitudes and expectations towards computerised self-help treatments for mental health problems. *BMC Psychiatry* 2014 Apr 11;14:109 [FREE Full text] [doi: [10.1186/1471-244X-14-109](https://doi.org/10.1186/1471-244X-14-109)] [Medline: [24725765](https://pubmed.ncbi.nlm.nih.gov/24725765/)]
19. Schulze N, Reuter SC, Kuchler I, Reinke B, Hinkelmann L, Stöckigt S, et al. Differences in Attitudes Toward Online Interventions in Psychiatry and Psychotherapy Between Health Care Professionals and Nonprofessionals: A Survey. *Telemed J E Health* 2019 Oct 09;25(10):926-932. [doi: [10.1089/tmj.2018.0225](https://doi.org/10.1089/tmj.2018.0225)] [Medline: [30412450](https://pubmed.ncbi.nlm.nih.gov/30412450/)]
20. De Rosis S, Barsanti S. Patient satisfaction, e-health and the evolution of the patient-general practitioner relationship: Evidence from an Italian survey. *Health Policy* 2016 Nov;120(11):1279-1292 [FREE Full text] [doi: [10.1016/j.healthpol.2016.09.012](https://doi.org/10.1016/j.healthpol.2016.09.012)] [Medline: [27836231](https://pubmed.ncbi.nlm.nih.gov/27836231/)]
21. Insel TR. Digital Phenotyping: Technology for a New Science of Behavior. *JAMA* 2017 Oct 03;318(13):1215-1216. [doi: [10.1001/jama.2017.11295](https://doi.org/10.1001/jama.2017.11295)] [Medline: [28973224](https://pubmed.ncbi.nlm.nih.gov/28973224/)]
22. Barnett I, Torous J, Staples P, Sandoval L, Keshavan M, Onnela J. Relapse prediction in schizophrenia through digital phenotyping: a pilot study. *Neuropsychopharmacology* 2018 Jul;43(8):1660-1666 [FREE Full text] [doi: [10.1038/s41386-018-0030-z](https://doi.org/10.1038/s41386-018-0030-z)] [Medline: [29511333](https://pubmed.ncbi.nlm.nih.gov/29511333/)]
23. Bernardos A, Pires M, Ollé D, Casar J. Digital Phenotyping as a Tool for Personalized Mental Healthcare. In: *PervasiveHealth'19: Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*. 2019 Presented at: 13th EAI International Conference on Pervasive Computing Technologies for Healthcare; May 2019; New York p. 403-408. [doi: [10.1145/3329189.3329240](https://doi.org/10.1145/3329189.3329240)]
24. Topp CW, Østergaard SD, Søndergaard S, Bech P. The WHO-5 Well-Being Index: a systematic review of the literature. *Psychother Psychosom* 2015;84(3):167-176 [FREE Full text] [doi: [10.1159/000376585](https://doi.org/10.1159/000376585)] [Medline: [25831962](https://pubmed.ncbi.nlm.nih.gov/25831962/)]
25. Sánchez-López MP, Dresch V. The 12-Item General Health Questionnaire (GHQ-12): reliability, external validity and factor structure in the Spanish population. *Psicothema* 2008 Nov;20(4):839-843. [Medline: [18940092](https://pubmed.ncbi.nlm.nih.gov/18940092/)]
26. World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*. Geneva: World Health Organization; 2003.
27. Busner J, Targum SD. The clinical global impressions scale: applying a research tool in clinical practice. *Psychiatry (Edgmont)* 2007 Jul;4(7):28-37 [FREE Full text] [Medline: [20526405](https://pubmed.ncbi.nlm.nih.gov/20526405/)]
28. Alwin DF. Feeling Thermometers Versus 7-Point Scales. *Sociological Methods & Research* 2016 Jun 30;25(3):318-340. [doi: [10.1177/0049124197025003003](https://doi.org/10.1177/0049124197025003003)]
29. Kieruj ND, Moors G. Variations in Response Style Behavior by Response Scale Format in Attitude Research. *International Journal of Public Opinion Research* 2010 Jul 23;22(3):320-342. [doi: [10.1093/ijpor/edq001](https://doi.org/10.1093/ijpor/edq001)]
30. Griffiths TL, Ghahramani Z. The Indian buffet process: An introduction and review. *J Mach Learn Res* 2011;12:1185-1224 [FREE Full text]
31. Pradier MF, Stojkoski V, Utkovski Z, Kocorev L, Perez-Cruz F. Sparse Three-Parameter Restricted Indian Buffet Process for Understanding International Trade. 2018 Presented at: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); April 15-20, 2018; Calgary, AB, Canada. [doi: [10.1109/ICASSP.2018.8462666](https://doi.org/10.1109/ICASSP.2018.8462666)]
32. Mehrotra K, Mohan C, Ranka S. *Elements of Artificial Neural Networks*. Cambridge, Massachusetts: MIT Press; 1996.
33. Barrigón ML, Berrouguet S, Carballo JJ, Bonal-Giménez C, Fernández-Navarro P, Pfang B, MEMind study group. User profiles of an electronic mental health tool for ecological momentary assessment: MEMind. *Int J Methods Psychiatr Res* 2017 Mar;26(1) [FREE Full text] [doi: [10.1002/mpr.1554](https://doi.org/10.1002/mpr.1554)] [Medline: [28276176](https://pubmed.ncbi.nlm.nih.gov/28276176/)]
34. Gómez-Carrillo A, Barrigón M, Leon-Velasco A, Gonzalez-Garrido C, Ruiz-Gomez M, Molina-Madueño R, et al. How far is clinical assessment from the bullseye? Using MEMind to compare clinical assessment with self-assessment in patients with depression and anxiety diagnosis. *The European Journal of Psychiatry* 2017 Oct;31(4):158-164. [doi: [10.1016/j.ejpsy.2017.08.006](https://doi.org/10.1016/j.ejpsy.2017.08.006)]
35. Berrouguet S, Baca-García E, Brandt S, Walter M, Courtet P. Fundamentals for Future Mobile-Health (mHealth): A Systematic Review of Mobile Phone and Web-Based Text Messaging in Mental Health. *J Med Internet Res* 2016 Jun 10;18(6):e135 [FREE Full text] [doi: [10.2196/jmir.5066](https://doi.org/10.2196/jmir.5066)] [Medline: [27287668](https://pubmed.ncbi.nlm.nih.gov/27287668/)]
36. Bakker D, Rickard N. Engagement in mobile phone app for self-monitoring of emotional wellbeing predicts changes in mental health: MoodPrism. *J Affect Disord* 2018 Feb;227:432-442. [doi: [10.1016/j.jad.2017.11.016](https://doi.org/10.1016/j.jad.2017.11.016)] [Medline: [29154165](https://pubmed.ncbi.nlm.nih.gov/29154165/)]
37. DuBenske LL, Gustafson DH, Namkoong K, Hawkins RP, Atwood AK, Brown RL, et al. CHES improves cancer caregivers' burden and mood: results of an eHealth RCT. *Health Psychol* 2014 Oct;33(10):1261-1272 [FREE Full text] [doi: [10.1037/a0034216](https://doi.org/10.1037/a0034216)] [Medline: [24245838](https://pubmed.ncbi.nlm.nih.gov/24245838/)]
38. Mangin D, Parascandolo J, Khudoyarova O, Agarwal G, Bismah V, Orr S. Multimorbidity, eHealth and implications for equity: a cross-sectional survey of patient perspectives on eHealth. *BMJ Open* 2019 Feb 12;9(2):e023731 [FREE Full text] [doi: [10.1136/bmjopen-2018-023731](https://doi.org/10.1136/bmjopen-2018-023731)] [Medline: [30760515](https://pubmed.ncbi.nlm.nih.gov/30760515/)]
39. Wernhart A, Gahbauer S, Haluza D. eHealth and telemedicine: Practices and beliefs among healthcare professionals and medical students at a medical university. *PLoS One* 2019;14(2):e0213067 [FREE Full text] [doi: [10.1371/journal.pone.0213067](https://doi.org/10.1371/journal.pone.0213067)] [Medline: [30818348](https://pubmed.ncbi.nlm.nih.gov/30818348/)]
40. Raychaudhuri K, Ray P. Privacy Challenges in the Use of eHealth Systems for Public Health Management. *International Journal of E-Health and Medical Communications* 2010;1:12-23. [doi: [10.4018/jehmc.2010040102](https://doi.org/10.4018/jehmc.2010040102)]

41. Ware P, Bartlett SJ, Paré G, Symeonidis I, Tannenbaum C, Bartlett G, et al. Using eHealth Technologies: Interests, Preferences, and Concerns of Older Adults. *Interact J Med Res* 2017 Mar 23;6(1):e3 [FREE Full text] [doi: [10.2196/jmir.4447](https://doi.org/10.2196/jmir.4447)] [Medline: [28336506](https://pubmed.ncbi.nlm.nih.gov/28336506/)]
42. Bucci S, Barrowclough C, Ainsworth J, Machin M, Morris R, Berry K, et al. Actissist: Proof-of-Concept Trial of a Theory-Driven Digital Intervention for Psychosis. *Schizophr Bull* 2018 Aug 20;44(5):1070-1080 [FREE Full text] [doi: [10.1093/schbul/sby032](https://doi.org/10.1093/schbul/sby032)] [Medline: [29566206](https://pubmed.ncbi.nlm.nih.gov/29566206/)]
43. Anderson REE, Spence SH, Donovan CL, March S, Prosser S, Kenardy J. Working alliance in online cognitive behavior therapy for anxiety disorders in youth: comparison with clinic delivery and its role in predicting outcome. *J Med Internet Res* 2012 Jun 28;14(3):e88 [FREE Full text] [doi: [10.2196/jmir.1848](https://doi.org/10.2196/jmir.1848)] [Medline: [22789657](https://pubmed.ncbi.nlm.nih.gov/22789657/)]

## Abbreviations

**CGI:** clinical global impression

**IBP:** Indian Buffet Process

**ICD:** International Statistical Classification of Diseases and Related Health Problems

**SPFM:** Sparse Poisson Factorization Model

*Edited by J Torous; submitted 20.11.19; peer-reviewed by S Goodday, A Daros; comments to author 10.01.20; revised version received 31.03.20; accepted 05.04.20; published 20.01.21*

*Please cite as:*

*Lopez-Castroman J, Abad-Tortosa D, Cobo Aguilera A, Courtet P, Barrigón ML, Artés A, Baca-García E*

*Psychiatric Profiles of eHealth Users Evaluated Using Data Mining Techniques: Cohort Study*

*JMIR Ment Health* 2021;8(1):e17116

URL: <http://mental.jmir.org/2021/1/e17116/>

doi: [10.2196/17116](https://doi.org/10.2196/17116)

PMID: [33470943](https://pubmed.ncbi.nlm.nih.gov/33470943/)

©Jorge Lopez-Castroman, Diana Abad-Tortosa, Aurora Cobo Aguilera, Philippe Courtet, Maria Luisa Barrigón, Antonio Artés, Enrique Baca-García. Originally published in *JMIR Mental Health* (<http://mental.jmir.org>), 20.01.2021. 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 Mental Health*, is properly cited. The complete bibliographic information, a link to the original publication on <http://mental.jmir.org/>, as well as this copyright and license information must be included.