Trackable Me: Relevant Data and User Types for the Tracking of Patient-Generated Health Data in Depression Care

Philipp Reindl-Spanner
Technical University of Munich
philipp.spanner@tum.de

Barbara Prommegger
Technical University of Munich
barbara.prommegger@tum.de

Jochen Gensichen
Hospital of the Ludwig-Maximilians-University
jochen.gensichen@med.uni-muenchen.de

Helmut Krcmar
Technical University of Munich
helmut.krcmar@tum.de

Abstract

Using patient-generated health data (PGHD) in depression care can provide valuable insights into patients' health. Due to the technical possibilities, patients can collect many PGHD types. However, these are not necessarily highly relevant to depression and are not considered relevant by all users. We, therefore, examined the relevance of various PGHD types for the treatment of depression and identified different types of users based on their data preferences. We surveyed 170 participants with depression and created a ranking for the most relevant data types. With subsequent cluster analysis, we identified four different user types: "Track-it-alls", "Medical Trackers", "Psychological Trackers," and "Untrackables". Based on these clusters, we show different possibilities for which user group and which types of PGHD are most suitable. With the results of this paper, we underline the need for tailored PGHD apps to improve personalized care in depression treatment.

Keywords: Health information technology (HIT), Mental Health Care, Patient-Centered Care, Mobile Applications

1. Introduction

During the COVID-19 pandemic, mental illness grew substantially in global populations. For example, the prevalence of depression in the general population globally increased to approximately 25%, compared to a worldwide prevalence of about 3.44% in 2017 (Bueno-Notivol et al., 2021). Along with the significant rise in cases, Aziz et al. (2022) report a significant increase in mental health app users in 2020 (during COVID-19) compared to 2019 (before COVID-19). These mental health apps, easily accessible on smartphones (Torous et al., 2018), assist in monitoring and treating patients' symptoms (Nittas et al., 2019), making them valuable tools for supporting public mental health. The most common disease treated through these apps is depression (Larsen et al., 2016).

Within this category of mental health apps, apps that use patient-generated health data (PGHD) stand out. Unlike general mental health apps, PGHD apps collect and use data generated by patients themselves (Shapiro et al., 2012). PGHD refers to patients' health-related information produced, documented, and gathered by patients. It encompasses a wide range of data, including the patient's health and treatment history, symptoms experienced, lifestyle choices, and other relevant details (Shapiro et al., 2012). The utilization of such data in the healthcare industry has been established for many years, for example, for treating diabetes or hypertension (Shah & Garg, 2015; Turner et al., 2021), and enable healthcare professionals to gain a deeper understanding of their patients' lives (Burgermaster et al., 2020). Its collection used to rely on paper documentation or specialized devices with hardware components (Cahn et al., 2018). However, the widespread adoption of smartphones and smart devices has brought about a revolutionary change in how PGHD is gathered nowadays and opens the usage of PGHD in new areas.

Due to the large number of different types of PGHD, such as sleep data, vital parameters, activity data, or data on smartphone usage, it is crucial to identify data types that provide relevant insights into the disease. It is, therefore, essential to reduce the number of PGHD types collected to reduce the effort patients and doctors spend collecting and evaluating the data. Using only relevant types of PGHD ensures that the apps remain clear and usable, and that relevant data is not overlooked (Jim et al., 2020).

In addition to identifying relevant types of data for depression care from the patient's perspective, it is

URI: https://hdl.handle.net/10125/109272
978-0-9981331-8-8

equally important to identify and understand the characteristics and preferences of the target group for the use of PGHD in depression care. This way, patients can be assured of PGHD's accessibility, and more users can be engaged in data tracking.

We employed a quantitative research approach and conducted a survey on the relevance and user characteristics of PGHD for depression care. Therefore, this paper aims to answer the following two research questions:

RQ1: Which types of data are considered relevant for the treatment of depression by patients with depression?

RQ2: What are the user types for the usage of PGHD in depression care?

The remaining article is structured as follows: First, we provide a theoretical background on personalized care and PGHD in depression care. We then describe the study design and the cluster analysis in detail. After that, we present the results, introducing four user types for tracking PGHD in depression care. Finally, we discuss the results and suggest ways our findings could be used to design PGHD apps.

2. Theoretical Background

The following section discusses personalized care and PGHD in depression care.

2.1 Personalized Care

Personalized care is a treatment method that focuses on the patient's unique needs, characteristics, and preferences. It involves collaboration between clinicians and patients and considers the care recipient's perspectives, experiences, and health-related data (Coulter et al., 2015). By tailoring healthcare to each individual, personalized care aims to make treatments more effective, improve patient satisfaction, and avoid unnecessary procedures. This approach shifts from the traditional method of treating everyone equally and instead seeks to provide more precise and individualized care. Personalized care is part of the broader concept of patient-centered care, emphasizing that healthcare systems must more effectively address patient needs (European Commission, 2020).

In the case of complex chronic diseases such as depression, the use of personalized care, which focuses on the patient, is essential to be able to take all aspects of the disease into account in therapy (Menear et al., 2022). Different theoretical care models, such as the Chronic Care Model, aim to treat patients better by considering their needs and wishes (Gensichen et al., 2022; Wagner et al., 1996).

Using PGHD in these approaches can provide further insights into the illnesses, offering the practitioner continuous objective and subjective impressions of the patient's life (Wu et al., 2020). Identifying preferences and user types plays an important role here so that the right types can be selected for the patient when using PGHD in personalized care.

2.2 Patient-Generated Health Data in Depression Care

The growing demand for mental health care necessitates more efficient and accurate approaches. Information and Communication Technology offers scalable solutions and enables a more comprehensive range of data collection through wearable sensors, mobile apps, and smart devices. These technologies allow for passive data collection of PGHD and a more precise diagnosis (Wang et al., 2018).

By definition, PGHD is health-related data created, recorded, and gathered. Integrating PGHD into healthcare workflows significantly benefits diagnosis and treatment (Shapiro et al., 2012). PGHD enables new avenues for diagnosing and treating medical conditions (Burgermaster et al., 2020). Furthermore, including PGHD leads to a transformative shift in consultation and treatment planning (Burns et al., 2019). By incorporating PGHD, healthcare providers can offer improved and personalized treatment approaches for various diseases (Cahn et al., 2018).

Various types of PGHD are suitable for the treatment of depression, not only to support the treating clinician during the therapy but also the patient in monitoring their disease. This data can be divided into different categories. Firstly, "physical health" data provides information about the patient's vital signs. This includes but is not limited to, the heart rate (Ng et al., 2019), blood pressure (Turner et al., 2021), or data on patients' medication (Park et al., 2019). The second group of PGHD for the care of depression is behavioral and lifestyle data. This includes, but is not limited to, patient activity (Kim et al., 2017) or data on sleep (sleep duration and quality) (Ng et al., 2019; Tsuno et al., 2005). Finally, other relevant data is not health-related but contextual and provides further insights into the patient's environment (Shapiro et al., 2012). This includes, for example, data on the weather (Brazienė et al., 2022).

However, not all data that the patient can collect is relevant to the diagnosis or treatment of their condition, and collecting PGHD can sometimes be burdensome for patients (e.g., paper mental health questionnaires) (Piras, 2019). In addition, these data collection methods can make patients more prone to data collection errors. However, less error-prone automated data collection methods often result in patients having privacy concerns (Ng et al., 2019).

Overall, integrating PGHD in healthcare workflows holds tremendous potential to revolutionize diagnosis, treatment planning, communication, and shared decision-making in therapy. PGHD presents numerous opportunities for integration into depression care, although research on this topic is still limited. One area that can be extended with the introduction of PGHD is data-based care, where reliable sources of data collection are needed to support established instruments like questionnaires (Fortney et al., 2018). PGHD can improve treatment decision-making by systematically tracking symptoms and patient progress. Other studies suggest that patients should collect PGHD themselves and discuss it during counseling sessions, as this has been shown to enhance therapeutic feedback (Meng et al., 2018). Furthermore, PGHD can be used as a predictor for depression, utilizing the collected data to generate comparative values based on validated questionnaires to predict clinical risks. This predictive capability can identify "at-risk" patients who may require clinical intervention, such as those who are not engaged in care or not responding to treatment (Hallgren et al., 2017; Saeb et al., 2015).

3. Method

To answer the research questions, we surveyed 170 patients with depression about their assessment of PGHD for depression treatment. As a result, we identified the most relevant types of PGHD and clustered the participants' PGHD preferences to identify four distinct user types for the use of PGHD in the context of depression. The following explains the study design, data collection, and analysis.

3.1 Study Design

With the theoretical background of personalized care, we decided to survey people with depression to find out the PGHD preferences of patients with depression and to identify possible user types. As we recruited through the online survey platform "Prolific", the potential participants were asked whether they had ever had depression in their lives. If the answer to this question was yes, we included the participants in the study. In this survey, we asked participants how relevant they thought certain types of PGHD were for the care of depression using a PGHD mental health app.

We collected the data for this article as part of a more extensive study that asked depression patients about the functionalities and aspects of mental health apps with PGHD usage. During the questionnaire development for this study, we worked closely with psychologists and psychiatrists to ensure that the structure of the questionnaire and questions were ethically justifiable and provided a safe environment for the participants.

For this survey, we selected different types of PGHD for depression care and asked the study participants to rate them on a 7-point Likert scale from "Very irrelevant" to "Very relevant." We conducted rigorous literature research to identify common types of PGHD in mental health and depression care. Based on the results of this review, we decided on the 21 types of PGHD included in this study (Austin et al., 2020; Ng et al., 2019). To evaluate the PGHD types, we assigned the individual data types to one of three thematic categories. The categories consisted of PGHD on "physical health," "behavioral and lifestyle data" and "other data". Types of PGHD that could not be specifically assigned to the other two blocks were grouped under "other data".

In addition to collecting the data types, we asked the participants two questions about their opinions on the relevance of data privacy of PGHD apps, as studies have shown that privacy concerns are one of the biggest concerns of mental health app users (Torous et al., 2018). Please find the detailed instructions for the survey in Table 1.

3.2 Data Collection

For this study, we defined our target group as adults (older than 18 years) who had already experienced depression in their lives. We decided not to specify any further inclusion or exclusion criteria to obtain a representation of the general population. Table 2 provides an overview of the demographic characteristics of the sample.

We collected data using the online survey platform Prolific in February 2024. A total of 187 participants completed the survey. We then conducted a rigorous data cleaning process, eliminating incomplete surveys and surveys that were not filled out correctly (e.g., rushed through the survey, pattern in answering the questions). This reduced the number of questionnaires to 170.

3.3 Data Analysis

First, we created a ranking of the PGHD types to identify the most relevant types for depression care. To do this, we calculated the means and Std-deviations of

In your opinion, how relevant are the following features for a PGHD app?	
PRIV1	The protection of personal data
PRIV2	Informing the user regarding what and when health data is sent to their care provider
How relevant do you consider…	
PHYS	…the following data on physical health for the treatment of depression within PGHD apps?
BEHAV	…the following data on behavior and lifestyle for the treatment of depression within PGHD apps?
OTHER	… the following other data for the treatment of depression within PGHD apps?

Table 2: Survey questions on Data Privacy and types of PGHD

the 21 types of PGHD. The results provide relative insight into the participants' relevance assessment. With this first analysis, we were able to identify relevant types of PGHD for depression care.

Next, we conducted a clustering algorithm based on the PGHD type preferences, using the "factoextra" library in R. We followed the approach of Prommegger et al. (2024), which conducted a similar cluster algorithm in a different context.

We first normalized the survey results of the data types to eliminate redundant data and ensure that good-quality clusters are generated (Virmani et al., 2015). We then calculated the optimal number of clusters using the total within sum-of-square method, also known as the elbow method. This method concluded that our data had three or four distinct clusters. Based on the assumption that three or four clusters are present, we performed ten runs for each number of clusters. We found that the within-cluster sum of distances decreases slowly after four clusters, indicating that four clusters are present. After deciding on four clusters, we conducted ANOVA analyses on the participants' demographics to identify significant differences in the distinct characteristics of the individuals within the clusters. In the following section, we will present the results of our analyses.

4. Results

In this section, we present our findings on the relevance of PGHD types, and the user types associated with the use of PGHD for depression care.

4.1 Ranking of Data Types

Table 3 presents the relative importance of the provided PGHD (based on mean), assessed by the participants. The five most relevant PGHD types were medication (mean=6.29), the measurement of suicidality (mean=6.20), somatic symptoms tracking

		Cluster 1, N = 65	**Cluster 2**, N = 38	**Cluster 3**, N = 44	**Cluster 4**, N = 23	**p-value**
Gender	Female	34 (52.3%)	23 (60.5%)	24 (54.5%)	10 (43.4%)	0.219
	Male	29 (44.6%)	11 (29%)	18 (40.9%)	11 (47.8%)	
	Other	1 (1.5%)	1 (2.6%)	0 (0%)	0 (0%)	
	Prefer not to say	1	3	2	2	
Degree	Some secondary education (high school or below)	0 (0%)	4 (10.5%)	1 (2.3%)	0 (0 %)	0.031*
	High school graduate or equivalent	25 (38.4%)	17 (44.7%)	25 (58.8%)	9 (39.1%)	
	Bachelor's Degree	28 (43.1%)	14 (36.8%)	13 (29.5%)	9 (39.1%)	
	Master's Degree or higher	12 (18.4%)	4 (10.5%)	5 (11.4%)	5 (21.7%)	
Age Group	Below 35	29 (44.6%)	7 (18.4%)	12 (27.3%)	9 (39.1%)	0.015*
	35-55	30 (46.1%)	22 (57.8%)	25 (56.8%)	12 (52.2%)	
	55 and above	6 (9.2%)	9 (23.9%)	7 (15.9%)	2 (8.7%)	
Data Privacy	Mean values	6.72	6.59	6.74	5.85	<0.001***
	* p < 0.05; ** p < 0.01; *** p < 0.001					

Table 1: Cluster details and differences based on demographics

(mean=6.09), mood (mean=6.09), and stress (mean=6.08).

These highly rated data types were followed by sleep (mean=6.04), mental health questionnaires (mean=6.02), weight (mean= 5.97), alcohol (mean=5.90), and routine (mean=5.80), which showed average to high importance.

Next, the patients ranked blood pressure (mean=5.74), diet (mean=5.66), pulse, and oxygen saturation (mean=5.66 and 5.52), reflecting their importance for cardiovascular and respiratory health.

At the lower end of the ranking, we could find hygiene (mean=5.46), steps (mean=5.18), social interactions (mean=5.11), and sports activities (mean= 4.91). Digital use (mean=4.58), weather (mean=4.13), and communication (mean=4.04) were assessed as the least relevant PGHD types.

We generally observed a higher standard deviation among lower-ranked PGHD, indicating a higher diversity for assessing relevance among different patients.

4.2 Cluster Analysis

The standard deviation, especially among lower-ranked PGHD types, hints at differences in the patient's relevance assessment. To gain more insights into this effect, we carried out a cluster analysis to identify different user types for using PGHD in the treatment of depression. In this section, we present the four identified user types for PGHD.

Table 2 and Table 3 provide an overview of identified user types. The first cluster (n=65; 38.2%) consists of people who find PGHD generally relevant for the treatment of depression. The second cluster (n=38; 22.4%) consists of people who find mainly psychology-related types of PGHD relevant. The third cluster (n=44; 25.9%) consists of people who find data on physical health particularly relevant. Finally, the fourth cluster (n=23; 13.5%) is generally averse to PGHD and finds hardly any type relevant for the treatment of depression. We found significant differences between these clusters concerning age groups, education, and their attitude toward data privacy. Table 2 provides an overview of the demographic characteristics of our sample and the individual clusters. We focus primarily on the gender, education, and age group of the participants and show the differences between the clusters and whether these are significant based on an ANOVA analysis. In the following, we will explain the individual clusters in detail.

Cluster 1, "Track it Alls", is defined by people who find all types of PGHD equally valuable for the care of depression. Individuals in this cluster rated almost all types of PGHD higher than the other clusters. These PGHD include sleep, daily steps, and sporting activity, which are easy to track and are often already tracked in the lifestyle area. This suggests that the individuals in this cluster are interested in tracking PGHD in the context of treating depression or even tracking data about themselves. Simultaneously, these results suggest that the individuals in this cluster may not distinguish which specific data types are most beneficial for depression treatment based on their generally very high ratings for all included data types.

Looking at Table 2, we can see that this cluster is the youngest and mainly consists of individuals with a university degree. For these reasons, it can be assumed that the individuals in the cluster were already familiar with PGHD. Due to the relatively young population in this cluster, it can be assumed that the individuals are primarily familiar with digital tracking forms.

The individuals in this cluster also rated the relevance of data protection and privacy of their data higher. This suggests that the individuals in this cluster understand data privacy concerning the use of PGHD and are, therefore, familiar with the data privacy issues associated with PGHD.

Based on their rankings and increased willingness to use wearables, we suggest this cluster is more accessible to enthuse about tracking data in a medical context and is well suited for additional treatment of their illness using PGHD.

Cluster 2, "Physical Trackers", differs from the first cluster in that the individuals in this cluster find a selection of PGHD, namely physical health data, most beneficial for treating depression. This data includes cardiac parameters, such as pulse or blood pressure, data on steps and sports activities, and other data, such as patients' symptoms, disease progression, and weight. In contrast to the first cluster, the individuals in this cluster tend not to rate most other data types as relevant, implying that they only associate high importance with physical factors. This suggests that the individuals in this cluster are aware of the physical component of depression. Cardiac problems are often associated with depression (Carney & Freedland, 2017), and monitoring other physical data, such as weight, can provide insights into the physical health of patients. Finally, correlations have also been found between patients' physical activity and depression (Paluska & Schwenk, 2000).

The average age of the individuals in this cluster is 44.6, the highest of all clusters. The combination of increased age and a preference for tracking physical health data suggests that these individuals may already have experience with physical illness or somatoform comorbidities of depression. The high rating of symptom and medication tracking (third highest overall, but still high) suggests that the individuals in this cluster are or have been undergoing medical treatment and are, therefore, aware of the importance of this data.

Cluster 3, "Psychological Trackers", is similar to Cluster 2 in that it only considers a selection of PGHD types relevant. However, respondents in Cluster 3 consider data on physical health less relevant but prefer data that can be directly linked to mental health. Therefore, this cluster's highly rated psychological PGHD, such as mental health questionnaires, mood, stress levels, and suicidal tendencies, are particularly relevant. It is not directly related to psychological data but is still highly relevant. Furthermore, this cluster rated medication tracking very relevant. In addition, individuals from Cluster 3 also value data that allows conclusions to be drawn about depression, such as data on sleep or data on substance abuse (e.g., alcohol or drug abuse) and therapy-relevant data like daily routine, hygiene, and social contacts. These ratings indicate that the individuals in this cluster know the importance of the direct collection of depression markers. The assessment of this cluster on psychologically and therapy-relevant data, especially the high rating of mental health questionnaires and the high rating for medication tracking, suggest that the individuals in this cluster have already received psychological treatment and are, therefore, familiar with these data types. This cluster also rated the relevance of data privacy the highest. This can be understood insofar as mental health data is highly sensitive data.

Cluster 4, "Untrackables", differs from the other three clusters. Compared to the other clusters, individuals in this cluster rated all types of PGHD as less relevant to depression care. These low scores may indicate that this cluster is generally less willing to collect data about themselves for depression care. There may be various reasons for this. Firstly, individuals may have less knowledge about PGHD or may indicate a general disinterest in collecting and providing this data. Equally, concerns about the accuracy or reliability of data collection may indicate a low data rating. It is also conceivable that this cluster prefers traditional treatment without PGHD or is skeptical about the use of PGHD in the treatment of depression. Finally, possible preferences of this cluster could also have been concealed by the pre-selection of the 21 types of PGHD for the survey.

In addition, this cluster also rated the need for data protection and privacy as the least relevant. This indicates that the primary reason for rejecting the PGHD presented is not primarily data privacy concerns. Regarding demographics, this cluster is just below average in terms of the average age and, therefore, tends to include younger people. In terms of education, individuals tend to have acquired a university degree.

5. Discussion

After having determined which types of data are considered most relevant for the treatment of depression by patients with depression and having introduced four user types for the usage of PGHD in depression care, we will now present our theoretical and practical contributions.

5.1 Theoretical Contribution

The ranking of data types underlines the relevance of PGHD for depression treatment. We presented respondents with 21 types of PGHD related to depression or mental health and had them ranked according to their relevance to depression. The ranking shows that the respondents rated the types of PGHD presented differently in terms of relevance. The types of PGHD ranked highest by the respondents may be the most promising for the treatment of depression from the perspective of depressed patients.

The highest-ranked data type, Medication Tracking, may have various benefits in treating depression. For example, medication tracking can help improve the patient's medication adherence (Park et al., 2019). The third most important data type, "Symptoms Tracking", can also be helpful here and indicate comorbidities and undesirable side effects of the medication.

The next type of PGHD, "suicidality", is of crucial importance in the context of depression. Therefore, respondents considered the recording of suicidality to be an essential data source. It is important to emphasize that measuring suicidality is particularly challenging (Franklin et al., 2017). Nevertheless, various metrics, especially in the context of psychological questionnaires, can be collected concerning suicidality (Kroenke et al., 2001).

The following highly rated types of data on mood, stress, and sleep have already been shown to be related to depression and are, therefore, well-established in questionnaires and data-tracking (Cummins, 2013; Sano et al., 2018; Tsuno et al., 2005).

In contrast to the types of data considered relevant by patients, it is also essential to consider which types of data are relevant for the treatment of depression from the perspective of physicians and psychiatrists. This is particularly important as healthcare providers often encounter irrelevant PGHD (West et al., 2018). By limiting the types of PGHD collected and shared with healthcare professionals, the accessibility and

			Cluster 1	Cluster 2	Cluster 3	Cluster 4
	Datatype	**Mean Overall**	**Track-it-Alls**	**Physical Trackers**	**Psychological Trackers**	**Untrackables**
	N (%)	170	65 (38.2%)	38(22.4%)	44(25.9%)	23(13.5%)
1	Medication	6.29	**6.66 (+0.37)**	6.37 (+0.08)	**6.41 (+0.12)**	4.91 (-1.38)
2	Suicidality	6.20	**6.46 (+0.26)**	5.71 (-0.49)	**6.59 (+0.39)**	5.52 (-0.68)
3	Symptoms	6.09	**6.52 (+0.43)**	**6.26 (+0.17)**	5.91 (-0.18)	4.90 (-1.19)
4	Mood	6.09	**6.57 (+0.48)**	5.37 (-0.72)	**6.43 (+0.34)**	5.26 (-0.83)
5	Stress	6.08	**6.63 (+0.59)**	5.60 (-0.44)	**6.18 (+0.14)**	5.13 (-0.91)
6	Sleep	6.04	**6.60 (+0.56)**	5.74 (-0.30)	**5.98 (-0.06)**	5.09 (-0.95)
7	Questionnaires	6.02	**6.43 (+0.46)**	5.50 (-0.47)	**6.32 (+0.35)**	5.17 (-0.80)
8	Weight	5.97	**6.50 (+0.53)**	**6.26 (+0.29)**	5.68 (-0.29)	4.52 (-1.45)
9	Substance Abuse	5.90	**6.46 (+0.63)**	5.63 (-0.20)	**6.27 (+0.44)**	4.04 (-1.79)
10	Daily Routine	5.80	**6.54 (+0.74)**	5.05 (-0.75)	**5.82 (+0.02)**	4.91 (-0.89)
11	Blood Pressure	5.74	**6.32 (+0.59)**	**6.18 (+0.45)**	4.91 (-0.82)	4.96 (-0.77)
12	Nutrition	5.66	**6.51 (+0.86)**	5.37 (-0.28)	**5.39 (-0.26)**	4.30 (-1.35)
13	Pulse	5.66	**6.29 (+0.67)**	**6.24 (+0.62)**	4.89 (-0.73)	4.52 (-1.10)
14	Oxygen Saturation	5.52	**6.22 (+0.66)**	**5.97 (+0.41)**	4.61 (-0.95)	4.57 (-0.99)
15	Hygiene	5.46	**6.32 (+1.06)**	4.11 (-1.15)	**5.73 (+0.47)**	4.74 (-0.52)
16	Daily Steps	5.18	**6.04 (+0.86)**	**5.13 (-0.05)**	4.34 (-0.84)	4.39 (-0.79)
17	Social	5.11	**6.05 (+0.94)**	3.79 (-1.32)	**5.32 (+0.21)**	4.22 (-0.89)
18	Sport	4.91	**6.00 (+1.09)**	**4.52 (-0.39)**	4.18 (-0.73)	3.87 (-1.04)
19	Digital Usage	4.58	**5.37 (+0.79)**	3.37 (-1.21)	**4.61 (+0.03)**	4.26 (-0.32)
20	Weather	4.13	**5.00 (+0.87)**	2.82 (-1.31)	**4.34 (+0.21)**	3.43 (-0.70)
21	Communication	4.04	**4.86 (+0.82)**	2.61 (-1.43)	**4.14 (+0.10)**	3.87 (-0.17)
	Functionalities					
	Data Privacy	6.58	**6.72 (+0.14)**	6.59 (+0.01)	**6.74 (+0.16)**	5.85 (-0.83)
		The two highest-ranked clusters are marked in bold.				

Table 3: Cluster breakdown of sample

usability of the integrated PGHD platform can be improved for treating professionals. This approach contributes to a better understanding of the relevant PGHD types in the treatment of depression and forms the basis for further research on the use of PGHD in these conditions.

We conclude that PGHD can be used to develop holistic and personalized healthcare solutions that incorporate valuable types of data to treat depression more effectively. These solutions can provide a more comprehensive treatment approach by considering medication, mood, physical activity, sleep, vital and cardiac signs, and psychological questionnaires.

Differences in needs for the personalization of health interventions: By identifying cluster-specific characteristics based on the relevance of different types of PGHD to depression treatment, we can show that these potential user groups have different requirements for tracking PGHD. For this reason, it is essential to tailor PGHD measures to both the disease and the needs of patients to increase their engagement in PGHD collection.

Cluster 1 had a general interest in PGHD for the treatment of depression. Individuals from this cluster can be recruited to collect various types of data. Therefore, this group is particularly well suited for clinicians and patients to work together to identify and collect the most relevant data types for the patient.

Cluster 2 had a high relevance score for data primarily reflecting physical characteristics. Although the individuals in this cluster are comparatively older than those in the other clusters, their preferences match well with the data types that can be collected by wearables (e.g., steps, sports activities, heart rate). Therefore, we suggest engaging these individuals in collecting PGHD through wearables.

Cluster 3 had an increased interest in data with psychological relevance. Most of the data categorized as relevant by this cluster is based on subjective perceptions. Studies have shown that the collection and integration of subjective paired with objective data can improve the understanding and treatment of depression. The combination of self-report (e.g., questionnaires) and passively collected data from sensors can provide deeper insights into the mental health status of patients (Nickels et al., 2021)At the same time, a high level of data privacy is essential for this cluster.

For clusters 2 and 3, it is imaginable to increase potential users' engagement by collecting a combination of the most relevant data types and the PGHD specifically rated highly by the two clusters.

Cluster 4 categorized all types of PGHD as less relevant. Several approaches are conceivable to convince the individuals in this cluster to track PGHD. Firstly, it is essential for this cluster that they are convinced of the relevance of PGHD in treating depression. Relevant PGHD that do not significantly impact people's everyday lives can then be used to make it easier for them to get involved in the PGHD collection. Furthermore, other motivational concepts can be used for this cluster. Therefore, it is possible to incorporate gamification elements such as achieving goals, completing challenges, or obtaining streaks into the apps to engage users in consistent data collection (Ilhan & Fietkiewicz, 2019).

5.2 Practical Contribution

Our practical implications are aimed at healthcare providers and technology developers looking to improve depression treatment through using PGHD.

From our theoretical findings, we can deduce that the cluster of **"Untrackables"** generally needs to be addressed differently than the other clusters, as this cluster generally does not see any relevance in PGHD for depression treatment. Therefore, to improve user engagement, app solutions developers must integrate educational components that communicate the importance of collecting different types of data, especially those data types that are considered less relevant overall.

Healthcare providers can develop more effective **targeted interventions**. Based on relevant types of PGHD, apps should offer immediate help or resources when suicidality measures indicate a high risk. This could include automated notifications to healthcare providers, contacts to crisis hotlines, or in-app counseling sessions. In addition, personalized medication adherence reminders and disease symptom tracking can be implemented to ensure users are continuously managing their conditions. These interventions can also include making patient recommendations based on real-time data applications. For example, if a user's mood or stress levels fluctuate wildly, the app could suggest relaxation techniques, cognitive behavioral therapy exercises, or connecting with a virtual therapist for immediate support.

Informed Clinical Decisions through integrating diverse PGHD enable clinicians to make more informed decisions about treatment plans. For example, understanding the interplay between medication adherence, sleep patterns, and mood fluctuations can help tailor interventions that address specific patient needs, leading to improved outcomes.

Health technology designers can use these behavioral insights we identified for the clusters to create more **user-centric products**. For instance, designing health apps and wearables that offer customizable data tracking options can personalize the app to the diverse needs of different user clusters, enhancing user engagement and satisfaction.

5.3 Limitations

Our study, like any other study, has limitations. First, we would like to discuss the context of the study. In this study, we focused on depressed patients living in the United States. Depression is a specific chronic illness, so our findings on PGHD are not necessarily transferable to other diseases and should be treated with caution when applied. The participants' origin, the cultural and educational characteristics of the US, and the prevalence of digital technologies there may also have influenced the results. Second, we believe that we may have limited participants in their choices by pre-selecting 21 types of PGHD for participants to evaluate. As a result, we may have missed data in the assessment that could have been uncovered through open-ended questions. Finally, in this study, we only consider the relevance of these types of PGHD from the patient's perspective. Although this perspective is critical, it is possible that healthcare professionals, such as psychiatrists or psychologists, have a different perspective on the types of data that are equally important but that we do not reflect in this study.

5.4 Future Research

Our study provides a reasonable basis for various future research projects. First, based on the used dataset, further insights can be obtained using covariance and codependency analysis methods. Furthermore, our study can be used as a basis for examining what influence the choice of the suitable types of PGHD has on the acceptance of mental health and depression apps. This can further explain how to tailor mental health apps to users to increase acceptance and effectiveness. This study offers the opportunity to be used as a baseline for how behavioral and lifestyle choices affect the choice and assessment of the relevance of PGHD for depression care. On this basis, strategies can be developed to integrate the types of PGHD that are not rated as relevant by patients but are still crucial for doctors, for example, into the survey without losing user acceptance. As noted in the Limitations, this study only considers the patient's perspective. Future studies can investigate healthcare professionals' preferences to determine which types of PGHD they find essential and whether there are clusters with different preferences within this target group.

6. Conclusion

Our study underlines the relevance of PGHD types in the treatment of depression and that different user types consider different types of PGHD to be relevant for the treatment of depression. Through our ranking of PGHD according to their relevance for depression treatment from the patient's perspective, we improve the understanding of patient preferences in the context of depression treatment using PGHD. The cluster analysis also reveals different user types, each with its own data needs and levels of engagement. The clusters showed differences primarily in age, education level, and attitudes toward the privacy of PGHD apps. These findings are crucial for developing user-centered mental health apps to improve patient engagement and treatment outcomes.

6. Acknowledgments

The research reported in this publication was funded by the German Research Foundation (DFG-GrK 2621/POKAL-Kolleg).

7. References

Austin, E., Lee, J. R., Amtmann, D., Bloch, R., Lawrence, S. O., McCall, D., Munson, S., & Lavallee, D. C. (2020). Use of patient-generated health data across healthcare settings: implications for health systems. *JAMIA open*, *3*(1), 70-76.

Aziz, M., Erbad, A., Almourad, M. B., Altuwairiqi, M., McAlaney, J., & Ali, R. (2022). Did usage of mental health apps change during COVID-19? A comparative study based on an objective recording of usage data and demographics. *Life*, *12*(8), 1266.

Brazienė, A., Venclovienė, J., Vaičiulis, V., Lukšienė, D., Tamošiūnas, A., Milvidaitė, I., Radišauskas, R., & Bobak, M. (2022). Relationship between depressive symptoms and weather conditions. *International journal of environmental research and public health*, *19*(9), 5069.

Bueno-Notivol, J., Gracia-García, P., Olaya, B., Lasheras, I., López-Antón, R., & Santabárbara, J. (2021). Prevalence of depression during the COVID-19 outbreak: A meta-analysis of community-based studies. *International journal of clinical and health psychology*, *21*(1), 100196.

Burgermaster, M., Son, J. H., Davidson, P. G., Smaldone, A. M., Kuperman, G., Feller, D. J., Burt, K. G., Levine, M. E., Albers, D. J., & Weng, C. (2020). A new approach to integrating patient-generated data with expert knowledge for personalized goal setting: A pilot study. *International journal of medical informatics*, *139*, 104158.

Burns, K., McBride, C. A., Patel, B., FitzGerald, G., Mathews, S., & Drennan, J. (2019). Creating consumer-generated health data: interviews and a pilot trial exploring how and why patients engage. *Journal of Medical Internet Research*, *21*(6), e12367.

Cahn, A., Akirov, A., & Raz, I. (2018). Digital health technology and diabetes management. *Journal of diabetes*, *10*(1), 10-17.

Carney, R. M., & Freedland, K. E. (2017). Depression and coronary heart disease. *Nature Reviews Cardiology*, *14*(3), 145-155.

Coulter, A., Entwistle, V. A., Eccles, A., Ryan, S., Shepperd, S., & Perera, R. (2015). Personalised care planning for adults with chronic or long-term health conditions. *Cochrane Database of Systematic Reviews*(3).

Cummins, R. A. (2013). Subjective well-being, homeostatically protected mood and depression: A synthesis. *The exploration of happiness: present and future perspectives*, 77-95.

European Commission. (2020). *Personalised medicine*. Retrieved June 6th from https://research-and-innovation.ec.europa.eu/research-area/health/personalised-medicine_en

Fortney, J. C., Unützer, J., Wrenn, G., Pyne, J. M., Smith, G. R., Schoenbaum, M., & Harbin, H. T. (2018). A tipping point for measurement-based care. *Focus*, *16*(3).

Franklin, J. C., Ribeiro, J. D., Fox, K. R., Bentley, K. H., Kleiman, E. M., Huang, X., Musacchio, K. M., Jaroszewski, A. C., Chang, B. P., & Nock, M. K. (2017). Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. *Psychological bulletin*, *143*(2), 187.

Gensichen, J., Lukaschek, K., Jung-Sievers, C., Falkai, P., Schmitt, A., Henningsen, P., Dreischulte, T., Pitschel-

Walz, G., Krcmar, H., & Böhm, M. (2022). Predictors and outcomes in primary depression care (POKAL)–a research training group develops an innovative approach to collaborative care. *BMC Primary Care*, *23*(1), 1-10.

Hallgren, K. A., Bauer, A. M., & Atkins, D. C. (2017). Digital technology and clinical decision making in depression treatment: Current findings and future opportunities. *Depression and Anxiety*, *34*(6), 494-501.

Ilhan, A., & Fietkiewicz, K. J. (2019). Learning for a healthier lifestyle through gamification: a case study of fitness tracker applications. *Perspectives on Wearable Enhanced Learning (WELL) Current Trends, Research, and Practice*, 333-364.

Jim, H. S., Hoogland, A. I., Brownstein, N. C., Barata, A., Dicker, A. P., Knoop, H., Gonzalez, B. D., Perkins, R., Rollison, D., & Gilbert, S. M. (2020). Innovations in research and clinical care using patient-generated health data. *CA: a cancer journal for clinicians*, *70*(3), 182-199.

Kim, Y., Heo, E., Lee, H., Ji, S., Choi, J., Kim, J.-W., Lee, J., & Yoo, S. (2017). Prescribing 10,000 steps like aspirin: designing a novel interface for data-driven medical consultations. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems,

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: validity of a brief depression severity measure. *Journal of general internal medicine*, *16*(9), 606-613.

Larsen, M. E., Nicholas, J., & Christensen, H. (2016). Quantifying app store dynamics: longitudinal tracking of mental health apps. *JMIR mHealth and uHealth*, *4*(3), e6020.

Menear, M., Girard, A., Dugas, M., Gervais, M., Gilbert, M., & Gagnon, M.-P. (2022). Personalized care planning and shared decision making in collaborative care programs for depression and anxiety disorders: a systematic review. *PLoS One*, *17*(6), e0268649.

Meng, J., Hussain, S. A., Mohr, D. C., Czerwinski, M., & Zhang, M. (2018). Exploring user needs for a mobile behavioral-sensing technology for depression management: qualitative study. *Journal of Medical Internet Research*, *20*(7), e10139.

Ng, A., Kornfield, R., Schueller, S. M., Zalta, A. K., Brennan, M., & Reddy, M. (2019). Provider perspectives on integrating sensor-captured patient-generated data in mental health care. *Proceedings of the ACM on human-computer interaction*, *3*(CSCW), 1-25.

Nickels, S., Edwards, M. D., Poole, S. F., Winter, D., Gronsbell, J., Rozenkrants, B., Miller, D. P., Fleck, M., McLean, A., & Peterson, B. (2021). Toward a mobile platform for real-world digital measurement of depression: user-centered design, data quality, and behavioral and clinical modeling. *JMIR Mental Health*, *8*(8), e27589.

Nittas, V., Lun, P., Ehrler, F., Puhan, M. A., & Mütsch, M. (2019). Electronic patient-generated health data to facilitate disease prevention and health promotion: scoping review. *Journal of Medical Internet Research*, *21*(10), e13320.

Paluska, S. A., & Schwenk, T. L. (2000). Physical activity and mental health. *Sports medicine*, *29*(3), 167-180.

Park, J. Y. E., Li, J., Howren, A., Tsao, N. W., & De Vera, M. (2019). Mobile phone apps targeting medication adherence: quality assessment and content analysis of user reviews. *JMIR mHealth and uHealth*, *7*(1), e11919.

Piras, E. M. (2019). Beyond self-tracking: Exploring and unpacking four emerging labels of patient data work. *Health Informatics Journal*, *25*(3), 598-607.

Saeb, S., Zhang, M., Kwasny, M., Karr, C. J., Kording, K., & Mohr, D. C. (2015). The relationship between clinical, momentary, and sensor-based assessment of depression. 2015 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth),

Sano, A., Taylor, S., McHill, A. W., Phillips, A. J., Barger, L. K., Klerman, E., & Picard, R. (2018). Identifying objective physiological markers and modifiable behaviors for self-reported stress and mental health status using wearable sensors and mobile phones: observational study. *Journal of Medical Internet Research*, *20*(6), e210.

Shah, V. N., & Garg, S. K. (2015). Managing diabetes in the digital age. *Clin Diabetes Endocrinol*, *1*, 16. https://www.ncbi.nlm.nih.gov/pubmed/28702234

Shapiro, M., Johnston, D., Wald, J., & Mon, D. (2012). *Patient-generated health data* RTI International, April,

Torous, J., Wisniewski, H., Liu, G., & Keshavan, M. (2018). Mental health mobile phone app usage, concerns, and benefits among psychiatric outpatients: comparative survey study. *JMIR Mental Health*, *5*(4), e11715.

Tsuno, N., Besset, A., & Ritchie, K. (2005). Sleep and depression. *Journal of clinical psychiatry*, *66*(10), 1254-1269.

Turner, K., Jo, A., Wei, G., Tabriz, A. A., Clary, A., & Jim, H. S. (2021). Sharing patient-generated data with healthcare providers: findings from a 2019 national survey. *Journal of the American Medical Informatics Association*, *28*(2), 371-376.

Virmani, D., Taneja, S., & Malhotra, G. (2015). Normalization based k means clustering algorithm. *arXiv preprint arXiv:1503.00900*.

Wagner, E. H., Austin, B. T., & Von Korff, M. (1996). Organizing care for patients with chronic illness. *The Milbank Quarterly*, 511-544.

Wang, K., Varma, D. S., & Prosperi, M. (2018). A systematic review of the effectiveness of mobile apps for monitoring and management of mental health symptoms or disorders. *Journal of psychiatric research*, *107*, 73-78.

West, P., Van Kleek, M., Giordano, R., Weal, M. J., & Shadbolt, N. (2018). Common barriers to the use of patient-generated data across clinical settings. proceedings of the 2018 CHI Conference on Human Factors in Computing Systems,

Wu, D. T., Xin, C., Bindhu, S., Xu, C., Sachdeva, J., Brown, J. L., & Jung, H. (2020). Clinician Perspectives and Design Implications in Using Patient-Generated Health Data to Improve Mental Health Practices: Mixed Methods Study. *JMIR Formative Research*, *4*(8), e18123.