

Closing the accessibility gap to mental health treatment with a conversational AI-enabled self-referral tool

Johanna Habicht^{1 *}, Sruthi Viswanathan¹, Ben Carrington¹, Tobias Hauser^{1,2}, Ross Harper¹, Max Rollwage¹

¹ Limbic Limited, London, UK

² Max Planck UCL Centre for Computational Psychiatry and Ageing Research
University College London, London WC1B 5EH, UK

* johanna@limbic.ai

Abstract

Accessing mental health care can be challenging, and minority groups often face additional barriers. This study investigates whether digital tools can enhance equality of access to mental health treatment. We evaluated a novel AI-enabled self-referral tool (a chatbot) designed to make entry to mental health treatment more accessible in a real-world setting.

In a multi-site observational study, data were collected from 129,400 patients who referred to 28 separate NHS Talking Therapies services across England. Our results indicate that the tool led to a 15% increase in total referrals, which was significantly larger than the 6% baseline increase observed in matched services using traditional self-referral methods during the same time period. Importantly, the tool was particularly effective for minority groups, which included non-binary (235% increase), bisexual (30% increase), and ethnic minority individuals (31% increase). This paints a promising picture for the use of AI chatbots in mental healthcare and suggests they may be especially beneficial for demographic groups that experience barriers to accessing treatment in the traditional care systems.

To better understand the reasons for this disproportional benefit for minority groups, we used thematic analysis and Natural Language Processing (NLP) models to evaluate qualitative feedback from 42,332 individuals who referred through the AI-enabled tool. We found that the tool's human-free nature and its ability to improve the perceived need for treatment were the main drivers for improved diversity.

These findings suggest that AI-enabled chatbots have the potential to increase accessibility to mental health services for all, and to alleviate barriers faced by disadvantaged populations. The results have important implications for healthcare policy, clinical practice, and technology development.

1 Introduction

Mental health is a global health priority, with the World Health Organization (WHO) recognising that mental health conditions are one of the leading causes of disability and disease burden worldwide (World Health Organisation, 2022). The most prevalent mental health disorders, such as anxiety disorders and depression, were estimated to affect 29% of the world’s population in their lifetime before the COVID-19 pandemic [Steel et al., 2014]. The pandemic further exacerbated the mental health crisis and increased the need for more support [Thome et al., 2021, Ornell et al., 2021, Loosen et al., 2021, Santomauro et al., 2021]. Besides the personal implications of mental health problems, anxiety disorders and depression alone lead to decreased productivity, causing the global economy to lose US \$1 trillion annually [Health, 2020]. By 2030, poor mental health is projected to cost \$6 trillion, underlining the urgent need to find solutions to address this global crisis [Health, 2020].

Mental health support can reduce the burden and has been found to be highly effective for many mental disorders. However, two main challenges preclude access to mental health support for many. Firstly, the supply of mental healthcare is not matching the demand as mental health services are widely underfunded and understaffed [World Health Organization, 2021b, Mahase, 2020]. Secondly, not every individual experiencing mental health problems seeks help, or they delay seeking it for years [Wang et al., 2007, Department of Health and Social Care, 2023], resulting in unmet needs and inadequate support at the right time. This can be due to barriers, such as a lack of perceived need for treatment, negative attitudes and stigma around mental health, as well as structural barriers [Andrade et al., 2014, Thornicroft et al., 2022]. These barriers can be even stronger for individuals among minority and disadvantaged backgrounds that may be stigmatised, such as individuals with ethnic or sexuality minority status [Clement et al., 2015, Thornicroft et al., 2022]. Digital technologies and Artificial Intelligence (AI) have been proposed as potential solutions to these challenges [Rudd and Beidas, 2020, Taylor et al., 2020, Kola et al., 2021]. While such technology has been shown to reduce the workload of the mental health staff and make services more efficient [Jayaraajan et al., 2022, Koutsouleris et al., 2022, Rollwage et al., 2022], less is known about the marginal impact of AI technologies in supporting individuals of differing demographics seeking help. Excitingly, digital solutions might be especially well-positioned to help individuals who are seeking support to overcome barriers, for example by offering more flexibility and reducing stigma [Rodríguez-Rivas et al., 2022].

The first step of many mental health care pathways is for individuals to seek help and be referred to the appropriate health care service. Ensuring easy and inclusive access at this initial step is crucial, regardless of an individual’s socioeconomic status, ethnicity, gender, or other such factors. This underpins a fair and equitable healthcare system. Unfortunately, evidence suggests that individuals from minority groups often face higher barriers and stigma in accessing care [Corrigan et al., 2014]. Referral to the appropriate healthcare service is a pivotal point in the treatment pathway, as failure to access the right support at the right time can lead to a worsening of symptoms, comorbidities, and adverse outcomes, including hospitalization or suicide [Williams et al., 2008, Reichert and Jacobs, 2018, De Girolamo et al., 2012]. Notably, a significant proportion of mental health problems remain undiagnosed and untreated, with estimates suggesting that as many as two-thirds of individuals with depression do not receive the necessary care [Moitra et al., 2022]. Therefore, making help-seeking more accessible is critical to providing timely and appropriate care to individuals in need, promoting better outcomes, and enhancing overall well-being.

The referral process for mental health care services varies between countries and services, and can involve self-referral or referral by a medical professional such as a general practitioner [Volpe et al., 2015]. The more efficient the referral process, the quicker support can be accessed, reducing the burden for the individual and the wider society. In the United Kingdom, the NHS Talking Therapies for Anxiety and Depression program - formerly Improving Access to Psychological Therapies (IAPT) - predominantly relies on self-referrals [Brown et al., 2010, The National Collaborating Centre for Mental Health, 2021]. Of the 1,740,652 patients referred to NHS Talking Therapies in 2021, 72% were self-referrals (NHS Digital). However, existing solutions may not be optimised or digitally streamlined, leading to lower completion rates and limited access to services [NHS Digital, 2019]. Thus, more optimal solutions are needed, as well as research into the evidence for these solutions.

To optimise this self-referral process, digital technologies have been suggested as a solution, such as AI chatbot solutions that have gained increased attention in healthcare [Tudor Car et al., 2020]. There is growing evidence that technology can improve help-seeking behaviour [Johnson et al., 2022] and that patients prefer interacting with chatbots for self-reported measures [Te Pas et al., 2020] and

discussing sensitive topics [Bhakta et al., 2014]. This suggests that digital technologies can remove some of the barriers to accessing mental health services, and thus increase accessibility.

To this end, we developed a novel AI-enabled chatbot solution for self-referrals, *Limbic Access*, which can optimise the standard referral process by autonomously gathering patient information to inform suitability for the service and initial presenting problem. We hypothesised that this AI-augmented self-referral could lower the threshold for accessing mental health services for three main reasons. Firstly, it can eliminate access barriers by providing a more convenient and flexible option than existing self-referral methods, such as via phone calls or static, daunting, and difficult-to-find webforms hosted on the service website. Additionally, the AI-enabled self-referral can be completed outside of service hours and with greater privacy. Secondly, a chatbot can guide patients proactively through the referral process, improving completion rates. Finally, chatbots have the potential to reduce the negative attitudes and stigma surrounding mental health and help-seeking. Therefore, we anticipated an increase in total referrals to services using the novel AI-enabled self-referral tool, as well as an increase in referrals from minority groups who face more barriers to access.

The use of AI chatbot solutions in healthcare is still an emerging area of research [Wilson and Marasoiu, 2022], with limited studies conducted in clinical settings. Therefore, our study aimed to assess the effectiveness of an AI-enabled self-referral tool for increasing accessibility to mental health services in a clinical context. To achieve this, we conducted an observational study using real-world data from the UK — a suitable environment for testing the self-referral tool due to its standardized referral process and a high percentage of self-referrals to NHS Talking Therapies.

Our results show that the AI-enabled self-referral tool, implemented in 14 services, had a positive impact on the number and diversity of referrals in a real-world setting. To further investigate the factors behind the improvement in diversity of access, we conducted thematic analysis on the qualitative feedback of 42,332 patients who used the tool, using Natural Language Processing (NLP) models. Our analysis showed that the tool’s ease of use and convenience were the primary drivers of the increase in referrals. Additionally, for minority groups, we identified two significant drivers: the tool’s human-free nature, which allowed patients to seek help without fear of being judged or discriminated against, and its ability to increase patients’ awareness of their treatment needs. These findings suggest that the tool may be able to alleviate the stigma experienced by minority groups when seeking mental healthcare.

2 Method

Here we studied the effects of an AI-enabled self-referral tool, *Limbic Access*, on the number and diversity of referrals to the NHS Talking Therapies for Anxiety and Depression (previously Improving Access to Psychological Therapies (IAPT) programme) in the United Kingdom. This tool has been developed with NHS Talking Therapies services and has been implemented as part of the routine mental healthcare in several NHS Talking Therapies services.

2.1 AI-enabled self-referral tool

The AI-enabled self-referral tool has been integrated into the service’s website (see Figure 1B) as a chatbot which is visible to any individual who visits the website. This reduces the barrier to seeking help as users do not need to manually search for how to get in contact with the service. The self-referral tool collects the necessary information for a referral required by the NHS Talking Therapies services (e.g. eligibility criteria, contact details, and demographic information). Additionally, the chatbot collects clinical information about the patient’s presenting symptoms, using questionnaires such as the Patient Health Questionnaire-9 (PHQ-9) [Kroenke et al., 2001], Generalised Anxiety Disorder Assessment (GAD-7) [Spitzer et al., 2006], the Work and Social Adjustment Scale (WSAS) [Mundt et al., 2002], and additional screening questions. These routine questionnaires and screening questions have not been typically collected at the point of referral in NHS Talking Therapies, but rather during the first assessment on a phone call. Frontloading the questionnaires during the referral process can reduce administrative burden, as well as make it easier for individuals to fill out these questionnaires in their own time and without having to speak to a person. Once the individual has completed the referral information using the chatbot, all the data is attached to the referral record within the NHS Talking Therapies services’ electronic health record to support the clinician in providing a high-quality and high-efficiency clinical assessment.

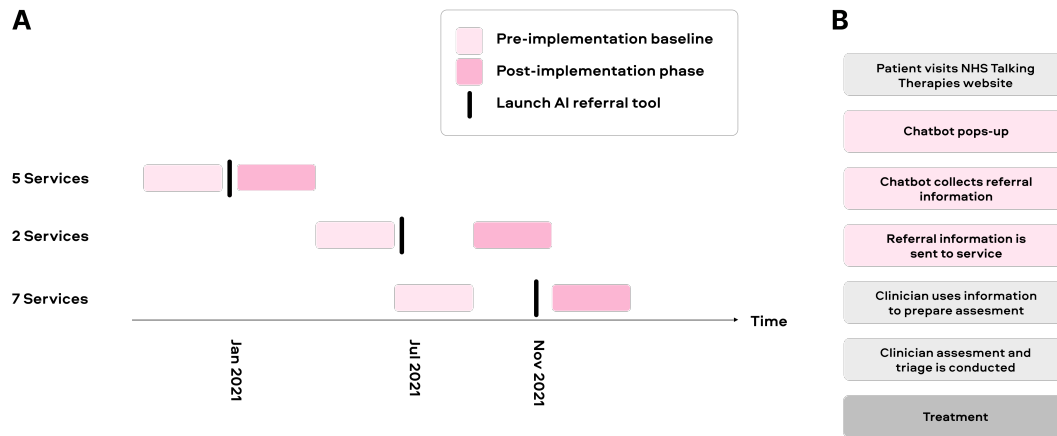


Figure 1: A) Study design for a real-world retrospective observational study with multiple sites, showing a three-month pre-implementation and three-month post-implementation period for different services. B) Referral workflow, showing the patient's journey from visiting the NHS Talking Therapies service's website to entering treatment.

2.2 Study design

This is a multi-site real-world retrospective observational study to investigate the impact of the AI-enabled self-referral tool on the number and diversity of referrals. The data used in this study are recorded in routine practice by all NHS Talking Therapies services and are reported to the public monthly, quarterly, and annually by NHS Digital ([NHS Digital](#)). The publicly available data undergoes quality checks and is a comprehensive dataset that facilitates high-quality analysis. The data used for the total number of referrals are taken from the annual reports and the demographic data is taken from the quarterly reports.

The analysis was conducted using data from ~129,400 patients who referred to 28 different NHS Talking Therapies services across England: 14 of these services implemented the AI-enabled self-referral tool and 14 of them did not while matching for the same timeframe. To understand the effects of the novel self-referral tool on the number and diversity of referrals, we investigated the time period before (pre-AI-enabled self-referral tool) and after the launch of the self-referral tool (post-AI-enabled self-referral tool) in the services. The timeframes for the pre- and post-periods were chosen as the closest quarter (3 months) to when the service launched the AI-enabled self-referral tool (see Figure 1A). Different services launched the self-referral tool at different times, allowing for a staggered study design to account for the fluctuation of referrals throughout the year due to seasonal and other factors.

To ensure that the effects observed were not confounded by other factors, we matched the services that implemented the AI-enabled self-referral tool to other similar services as commonly done in similar clinical and public health research studies [[Rose and Van der Laan, 2009](#)]. We only included control services with a webform as one of their ways of referral because this was the closest referral option to the AI-enabled self-referral tool investigated, thus representing an optimal control group. For each NHS Talking Therapies service using the AI-enabled self-referral tool, we calculated the total number of referrals in the pre- AI self-referral tool period. Then we calculated the total number of referrals for all other NHS Talking Therapies services during the same period of time. For example, when the tool was launched in July 2021, we calculated the number of total referrals for the second quarter (April to June) of 2021 for all services and used these metrics for matching the most similar control services. In order to match services, we calculated the Euclidean distance based on these two (normalized) features and took the service with the lowest distance to the NHS Talking Therapies service using the AI-enabled self-referral tool, where a low distance implies a high degree of similarity.

2.3 Analysis

2.3.1 Total number of referrals

To investigate whether the AI-enabled self-referral tool increased the number of referrals, we calculated the total number of referrals for each service in the pre- and post-implementation of the tool. To analyse the data between the pre- and post-implementation periods, we summed the total number of referrals for all services. To compare the change in the total number of referrals between services that implemented the AI-enabled self-referral tool and those that did not, we used a Chi-squared test.

To ensure that the total number of referrals did not increase due to other forms of referrals other than self-referrals (e.g. GP referrals) increasing in the time period, we repeated the analysis for only self-referrals in the same services at the same time period (Supplementary Figure 1).

2.3.2 Demographic data

To investigate whether the AI-enabled self-referral tool improved the diversity of referrals, we investigated the number of referrals for certain sociodemographic groups, focusing on gender identity, sexual orientation and ethnic groups. Specifically, we investigated the number of referrals for females, males, and non-binary individuals for gender identity; heterosexual, bisexual, and gay/lesbian individuals for sexual orientation; and Asian or Asian British, Black or Black British, Mixed, Other Ethnic groups, and White individuals for ethnicity. We calculated the total number of referrals from a certain sociodemographic group for the pre- and post-implementation periods and the percentage change for each group.

To further assess whether increased referrals from minority groups exceeded those of the most common groups, we constructed a logistic regression predicting on a single, inferred, patient level whether the patient referred pre- or post-implementation of the tool (0 = pre-implementation, 1 = post-implementation). As a predictor variable, we used the sociodemographic group of the individual, assigning the most common group (e.g. females for gender, heterosexuals for sexual orientation and white for an ethnic group) as the reference. We report the odds ratio, calculated as the exponential of the regression coefficient, along with confidence intervals.

2.3.3 NLP topic classification

To understand the mechanisms behind the increased access, we used thematic analysis powered NLP methods to analyse qualitative feedback given by patients who used the AI-enabled self-referral tool. NLP techniques, such as text classification, have gained popularity in healthcare settings for analysing free-text patient feedback to generate insight about patients' experience [Khanbhai et al., 2021]. In this study, we employed a supervised NLP classification model to categorise qualitative feedback provided by individuals upon completing their referrals via the AI-enabled self-referral tool to better understand the reasons for the increased number of referrals, especially from minority groups. From September 2021 to March 2023, a total of 157,416 patients used this novel self-referral approach, with 29% of these patients offering free-text feedback on their experience ($N = 46,166$). To ensure our analysis only included meaningful insights, we excluded any responses containing fewer than 10 characters, resulting in a final sample of 42,332 feedback entries. We employed a NLP classification model to categorize each feedback entry to a theme.

We first performed reflexive thematic analysis on feedback entries [Braun and Clarke, 2019], with one researcher open-coding a sample of 4,000 entries. Initial codes were discussed with the larger group of researchers, and a consensus was reached on the resulting themes (listed in the table below) after two meetings. Our team manually identified nine distinct themes, consisting of four positive themes and two neutral and three negative themes. We then selected entries under each theme to train a model on a subset of feedback entries (657 entries), ensuring that each theme had at least 50 representative entries. This subset was then utilised as the training and testing dataset for the NLP model.

Theme	Explanation
Convenient	Individuals mentioned that the referral process was easy and/or fast
Provided hope	Individuals mentioned that the referral was a hopeful first step towards getting support
Self-realisation	Individuals mentioned that the referral provided a self-realisation of their current situation, such as a realisation of suffering from a mental health condition or a need for treatment
Human-free	Individuals mentioned the human-free nature of the tool and removing the anxiety of talking to humans
Needed specific support	Individuals mentioned that they needed to talk about specific illnesses or other mental disorders
Other neutral feedback	Individuals mentioned any other neutral feedback
Expected support sooner	Individuals mentioned that they expected to start therapy faster
Wanted urgent support	Individuals mentioned that they felt that they needed immediate attention.
Other negative feedback	Individuals mentioned any other negative feedback

To prepare the feedback data for use with our model, we first used a sentence-BERT (SBERT) sentence transformer to transform each entry into numerical representations. We then employed Principal Component Analysis (PCA) to reduce dimensions. We used a multi-class logistic regression model and evaluated its accuracy through a 100-fold cross-validation approach. Our model achieved a micro-average F1 score of 0.64. This suggests that the model demonstrated a moderate level of performance in classifying patient feedback entries into their assigned themes.

To classify the 42,332 feedback entries into the nine themes, we employed the above-mentioned model. To determine whether the proportions of individuals mentioning specific themes differed between minority and majority groups, we used a Chi-squared test. To account for multiple comparisons, we used Bonferroni correction.

3 Results

3.1 Increase in the number and diversity of referrals

The implementation of the AI-enabled self-referral tool resulted in a 15% increase in the total number of referrals compared to a similar pre-implementation time period (Figure 2). In contrast, matched NHS Talking Therapies services with a similar number of total referrals in the pre-implementation period that used other self-referral methods, such as web forms, saw only a 6% increase in referrals in the same time period. The increase in total referral numbers for services using the AI-enabled self-referral tool was significantly higher than the increase for the matched control services ($\chi^2(1) = 86.3, p < 0.0001$). This suggests that the AI-enabled self-referral tool facilitates an increase in referrals compared to the national average in matched NHS Talking Therapy services during the same time period.

The results suggest that the AI-enabled self-referral tool has a generally positive impact on increasing access to mental health treatment for individuals. However, some disadvantaged sociodemographic groups may encounter more barriers when trying to access treatment. To explore this further, we investigated whether the tool's effect on increasing referrals holds true across multiple demographic groups or if it disproportionately benefits certain groups in accessing mental health care. We calculated the percentage change in the number of referrals from pre- to post-implementation of the AI-enabled self-referral tool, separately for all sociodemographic groups (see Supplementary Table 1 for the total

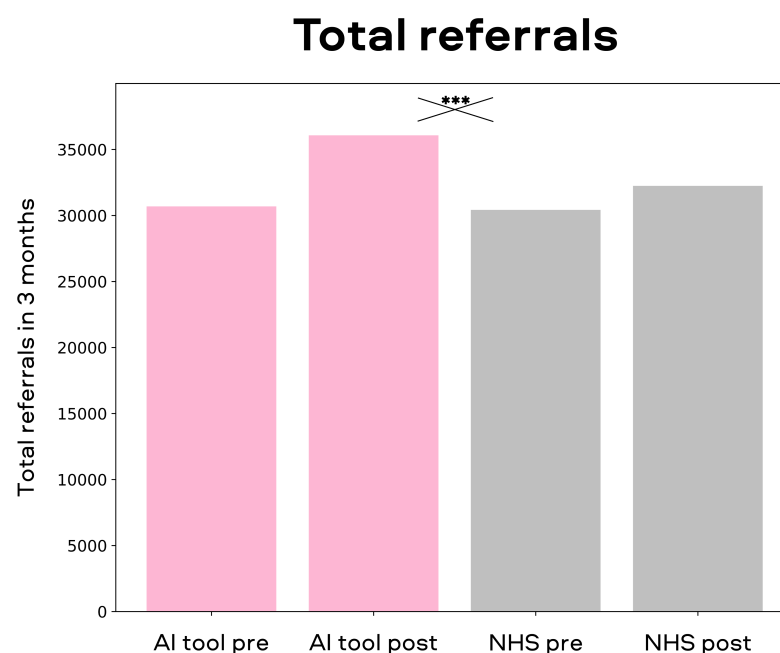


Figure 2: The total number of referrals pre- and post-implementation of the AI-enabled self-referral tool for services that implemented the novel AI tool (pink) and other similar services that did not implement the tool (grey). There was a 15% increase in total referrals for services that used the AI tool, whereas other services had a 6% increase in the same time period. The increase in referrals for services using the self-referral tool was significantly higher than the increase for other similar services. *** $p < 0.001$

number of referrals during pre- and post-implementation) and compared minority groups to the most common groups.

First, we focused on the gender identity of individuals who referred to the services. We found that referrals from individuals who identified as non-binary increased by 235%, whereas referrals from individuals who identified as female or male increased by 18% and 16%, respectively (Figure 3A). The increase in referrals from people who identify as non-binary was higher than females (odds-ratio = 2.83, CI = [2.264, 3.545], $p < 0.0001$) and males (odds-ratio = 2.90, CI = [2.314, 3.629], $p < 0.0001$), suggesting that the AI-enabled self-referral tool especially facilitates more access for people from gender minority groups.

Next, we investigated referrals based on individuals' sexuality. We found that referrals by people whose sexual orientation was bi-sexual increased by 30%, referrals by people whose sexual orientation is gay/lesbian increased by 19%, whereas referrals from heterosexuals increased by 14% (Figure 3B). The increase of people whose sexual orientation is bi-sexual was higher than for heterosexual patients (odds-ratio = 1.14, CI = [1.074, 1.219], $p < 0.0001$), suggesting that the tool facilitates more access for people who are not heterosexual. However, the increase in individuals identifying as homosexual did not significantly differ from the increase in referrals for heterosexual individuals (odds-ratio = 1.05, CI = [0.922, 1.189], $p = 0.216$).

Lastly, we evaluated the effects of the tool on referral rates for different ethnic groups. We identified a 31% increase in referrals for ethnic minorities, whereas the increase for the White group was 15% (Figure 3C), suggesting that different ethnic groups were more represented after the implementation of the AI-enabled self-referral tool. The increase in ethnic minorities was higher than for White individuals (odds-ratio = 1.14, CI = [1.081, 1.202], $p < 0.0001$), suggesting that the tool facilitates more access for people from ethnic minorities. To investigate these results further, we analysed ethnic groups in a more

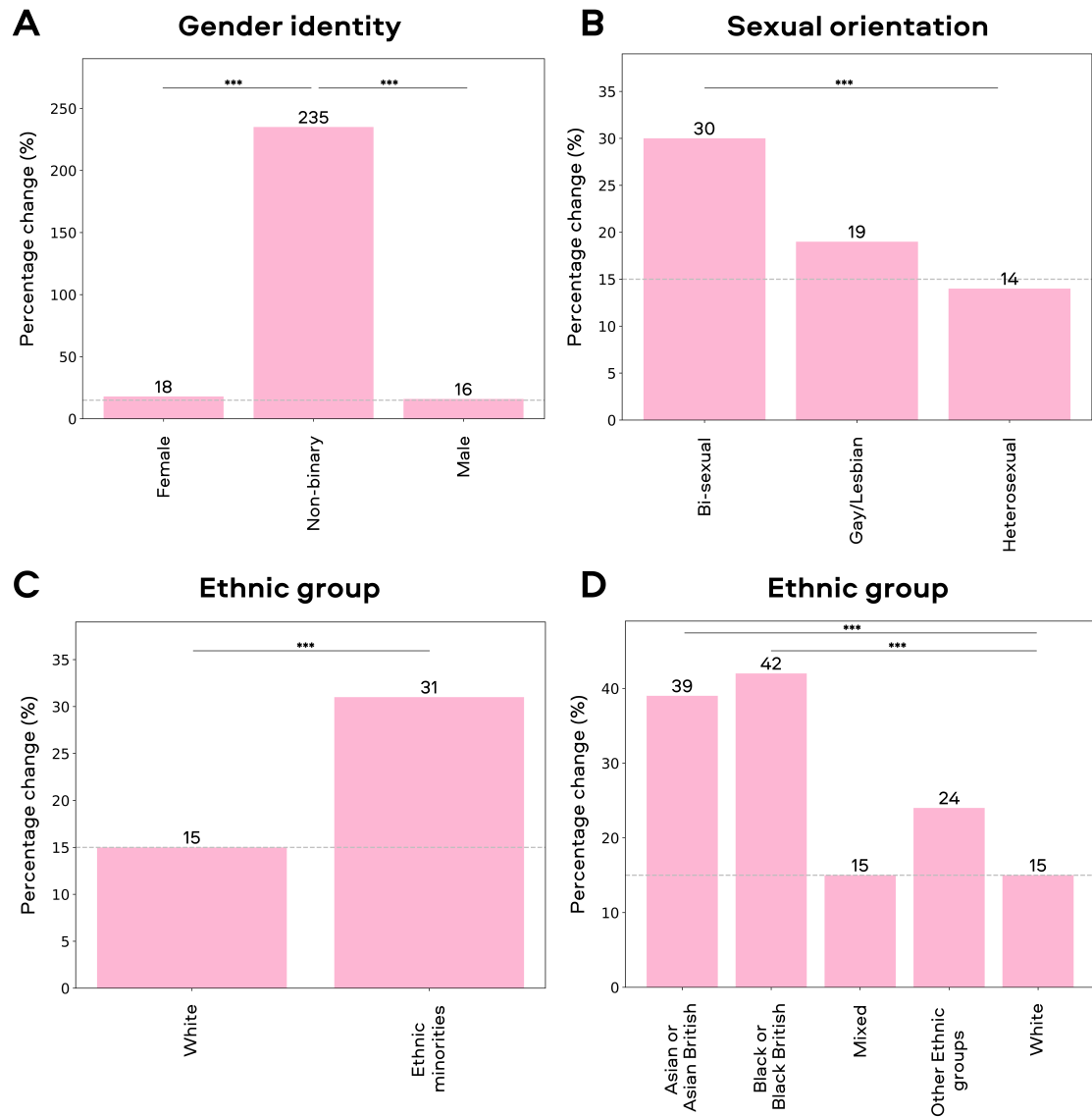


Figure 3: Percentage change for different sociodemographic groups for services that implemented the AI-enabled self-referral tool. The dotted line on each plot shows the overall percentage change in the total referral numbers (15%). A) A positive percentage change in referrals was observed for all genders, with the highest percentage increase in referrals for individuals who identify as non-binary. B) A positive percentage change in referrals was observed for people with all sexual orientations, with the highest percentage increase in referrals for individuals who are bi-sexual. C) A positive percentage change in referrals was observed for all ethnic groups, with the highest percentage increase in referrals for ethnic minorities. D) There was a higher positive change for Asian or Asian British and Black or Black British compared to the White group. *** $p < 0.001$

fine-grained manner. We particularly saw an increase for Asian and Asian British groups (odds-ratio = 1.21, CI = [1.116, 1.318], $p < 0.0001$) and Black and Black British groups (odds-ratio = 1.24, CI = [1.108, 1.378], $p < 0.0001$) compared to White individuals (Figure 3D). We did not see any differences in the number of referrals for individuals from other ethnic groups (odds-ratio = 1.08, CI = [0.913, 1.272], $p = 0.377$) and Mixed ethnic group (odds-ratio = 1.01, CI = [0.917, 1.107], $p = 0.877$) compared to White individuals.

In conclusion, these results suggest that the AI-enabled self-referral tool facilitates a general increase in referrals, as well as improves the diversity of referrals. An increase in referrals can potentially lead to longer wait times for clinical assessment and may reduce the number of people who access these assessments. To address this concern, we compared the percentage of people who were assessed in the time periods between the services that used the AI-enabled self-referral tool and services not using the tool (Supplementary Figure 3). A two-way ANOVA revealed no significant interaction between the time period and tool usage ($F(1, 13) = 0.09$, $p = 0.764$). Therefore, the AI-enabled self-referral tool did not negatively impact the number of people who accessed assessments.

3.2 Mechanism for the increase in referrals

To investigate the underlying mechanism behind the increased number and diversity of referrals through the AI-enabled self-referral tool, we analysed qualitative free-text feedback provided by 42,332 individuals upon completion of their referral. Our goal was to identify whether specific topics, such as specific advantages or disadvantages of the AI-enabled self-referral tool, were mentioned disproportionately by certain demographic groups. We used NLP topic classification to classify each feedback entry to a predefined theme (see Methods for details). We analysed feedback from individuals belonging to both minority and majority groups based on their ethnic background, sexual orientation, and gender identity. We compared minority groups where we observed differences in the number of referrals - bisexual individuals, non-binary and other gender minority individuals, and Asian, Black and other ethnic group individuals (Figure 4). We included the analysis of all minority groups in the Supplementary material (Supplementary Figure 2).

First, it is noteworthy that there are some generally overarching patterns across all demographic groups. Overall, 89% of the free text feedback was classified as positive, 7% neutral and 4% negative. In general, patients mentioned the convenient nature of the chatbot as the major advantage, emphasizing the effectiveness of this novel self-referral method. This was followed by the hope of having taken the first step to access support, the absence of human involvement reducing stigma and judgment and increased self-realization that they might suffer from a mental health condition and need support caused by the self-referral process. While these general patterns largely hold across all demographic groups, there were significant differences in the frequency with which these advantages were mentioned, indicating that different groups might experience different advantages which might contribute to the increased accessibility for these groups. The following sections highlight the significant findings, while additional results can be found in Supplementary Table 2-4.

We found that individuals from gender minority groups mentioned the absence of human involvement in the tool more frequently than females and males ($\chi^2(1) = 22.7$, $p < 0.0001$, Figure 4A). This suggests that the AI-enabled self-referral tool resolves an important barrier for this group by being human-free and indicates the mechanism for why we observed increased referrals from this minority group might be a reduction in stigma and judgment during this AI-enabled self-referral process. We observed a marginally significant reduction in the theme related to the self-referral providing hope for the gender minority groups ($\chi^2(1) = 8.84$, $p = 0.079$), which may suggest that the absence of human involvement in the tool is a more important factor for this group than the sense of hope it provides.

Similarly, we found that individuals who identify as bisexual mentioned the human-free nature of the tool more frequently than heterosexual individuals ($\chi^2(1) = 62.4$, $p < 0.0001$, Figure 4B) indicating that the human-free aspect is important for them. In addition, we saw a similar reduction in the theme about hope for bisexual individuals ($\chi^2(1) = 27.8$, $p < 0.0001$) indicating that the human-free aspect is more important for them compared to the feeling of hope provided by the tool.

When investigating ethnic minorities, we found that individuals from Asian and Black ethnic groups mentioned the self-realisation about the need for treatment theme more than White individuals ($\chi^2(1) = 46.6$, $p < 0.0001$, Figure 4C), which reflects that during the self-referral process, there was an increase in the individual's awareness that they might suffer from a mental health condition and the need for treatment. This may reflect how the tool can reduce the stigma faced by ethnic minority groups in

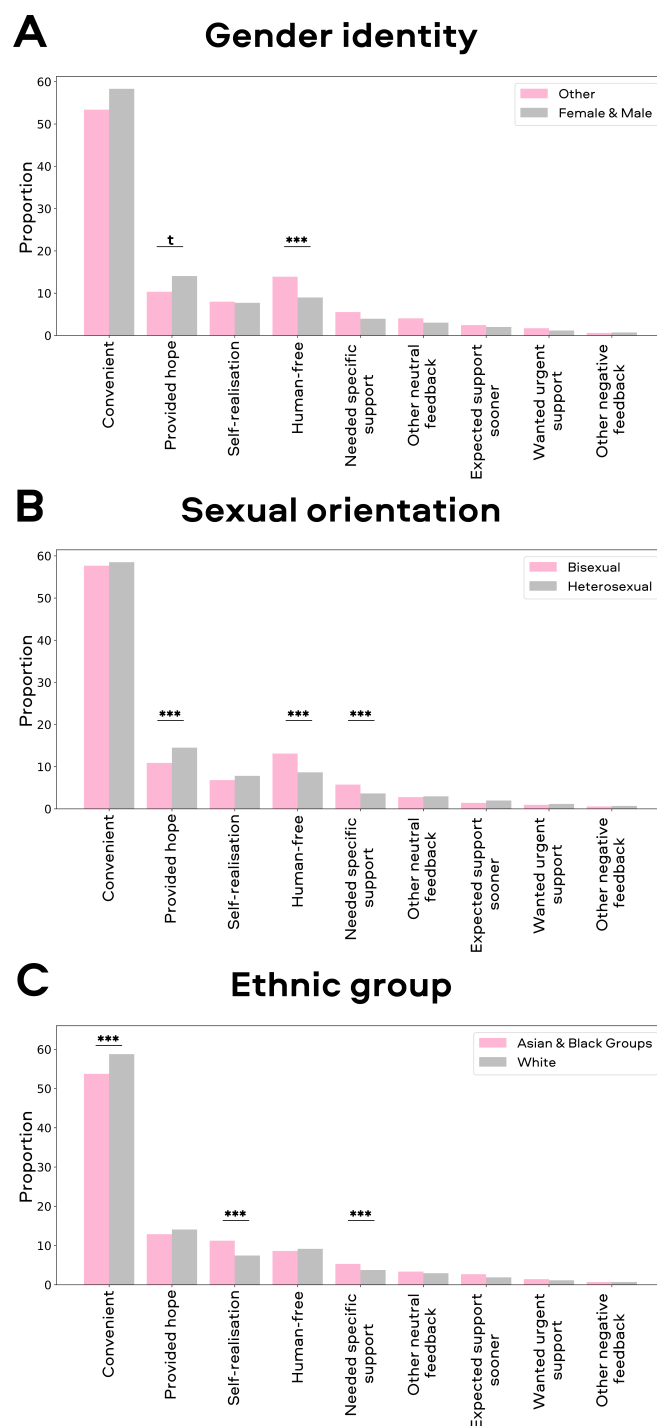


Figure 4: Breakdown of feedback themes from minority (pink bars) and majority (grey bars) groups. A) Individuals from gender minority groups mentioned the tool's human-free nature more compared to the majority groups. B) Bisexual individuals mentioned the tool's human-free nature more, as well as needing more specific support compared to the majority groups. C) Individuals from Asian and Black ethnic groups mentioned more the theme around self-realisation as well as needing more specific support compared to White individuals. *** $p < 0.001$; ** $p < 0.01$; t $p < 0.1$

seeking mental health support. In addition, we found that Asian and Black ethnic minority groups mentioned the convenient nature of the tool less than White individuals ($\chi^2(1) = 24.2, p < 0.0001$).

Finally, our analysis showed that individuals from Asian and Black ethnic groups and bisexual individuals provided more neutral feedback compared to majority groups, especially around the theme of needing specific support (sexuality: $\chi^2(1) = 30.8, p < 0.0001$; ethnicity: $\chi^2(1) = 14.6, p = 0.0036$). This could be attributed to the higher barriers faced by these individuals in seeking mental health support in general, as well as their specific problems related to their minority status, such as bisexual individuals seeking help for specific problems related to their sexuality.

In conclusion, our findings suggest that minority groups may face different barriers and value different aspects when seeking mental health treatment than majority groups. Some of these barriers might be especially well addressed by an AI-enabled self-referral tool's human-free aspect and the reduced stigma. Such solutions may remove these barriers and may be valued more by these minority groups, leading to a higher number of referrals.

4 Discussion

In this study, we investigated the effects of an AI-enabled self-referral tool on the number and diversity of referrals in the NHS Talking Therapies services in the United Kingdom. Notably, this is the first large-scale study using real-world data to investigate the potential impact of an AI-enabled tool on the quality of access to mental health care. We found that NHS Talking Therapies services that implemented the AI-enabled self-referral tool showed a 9 percentage point stronger increase in the number of referrals after the launch of this tool compared to matched NHS Talking Therapies services that used other forms of referral in the same time period. Specifically, services that used the AI-enabled self-referral tool showed a 15% increase in referrals, whereas the matched services that did not use the tool showed a 6% increase. This baseline increase is likely to reflect the national initiative to boost access across all NHS Talking Therapies services. At the time of writing, this self-referral tool has facilitated the referral of $\sim 157,000$ patients, which means that $\sim 14,000$ additional patients have found their way into treatment owing to this technical solution. Importantly, the increase in referrals did not negatively impact the number of individuals who accessed clinical assessments, which suggests that the tool did not exacerbate wait times for treatment. In addition, a separate study involving the same services demonstrated that this AI-enabled self-referral tool has further downstream implications and improves recovery rates [Rollwage et al., 2022].

An important insight from our study is that the services using the AI-enabled self-referral tool saw a proportionally higher increase in the number of individuals from minority groups, such as individuals who identify as non-binary, bisexual, and ethnic minorities. For individuals from ethnic minority backgrounds, we especially saw an increase for Asian or Asian British groups and Black or Black British groups. These groups have been previously underrepresented in mental healthcare treatment, indicating that this tool enabled a previously underserved need.

Findings from previous research suggest that digital technologies may facilitate help-seeking for mental health [Johnson et al., 2022] and make it easier for people to disclose personal information [Te Pas et al., 2020, Bhakta et al., 2014]. Our analysis of 42,332 patient feedback supports these findings, indicating that individuals are finding the new AI-enabled self-referral tool easy and convenient. This can be attributed to the tool's flexibility, as individuals can refer to the service outside of the usual hours and take the necessary time to complete the referral process. Furthermore, the AI-enabled self-referral tool automatically opens on the service's website, eliminating the need for patients to search for contact information or referral form. Another advantage is that individuals can more comfortably discuss their mental health issues without facing potential communication difficulties with a human. Addressing these barriers with an AI-enabled self-referral tool has likely lowered the threshold for seeking help, resulting in higher referral rates as indicated by our data.

Minority groups often face especially high barriers when seeking mental health treatment [Clement et al., 2015, Thornicroft et al., 2022], and our NLP analysis of patient feedback evaluated the mechanism of how the AI-enabled self-referral tool facilitated the increased referrals from minority groups. Interestingly, we found that minority groups mention different themes in their feedback, suggesting that the barriers they face and their needs are different to majority groups. We found that individuals from sexuality and gender minorities particularly valued the human-free nature of the self-referral tool, which enabled them to seek help without fear of judgment or discrimination. Ethnic minorities

mainly mentioned how the referral process helped them to realise that they might be suffering from mental health problems or the need for treatment which may reflect the negative attitudes and lack of perceived need for treatment these individuals might face. Furthermore, minority groups expressed a greater need for specific treatments, highlighting the importance of addressing their unique needs and barriers. These final points indicate that while these initial results are reassuring, further work needs to be done to address all barriers experienced by minority groups when accessing mental health treatment.

In interpreting the results of the study, it is important to consider certain limitations. One potential issue to consider is the possibility of cultural differences between various NHS Talking Therapies services. It is possible that those NHS Talking Therapies services making use of the AI-enabled self-referral tool are inherently more open to digital health approaches and innovation than those that have not adopted such tools, which could influence the referral numbers. Nevertheless, implementation of the AI-enabled self-referral tool was not part of wider digital innovation programs for any of the services and the increase in referrals was tightly linked to the launch of the AI-enabled self-referral tool.

Despite this limitation, our study also has noteworthy strengths. As the data was collected as part of a standard national data reporting process that undergoes quality checks, there is no concern about missing data or that the data is reported in different ways. In addition, the data was collected under real-world conditions from multiple services, which provides a larger sample size and thereby improves the generalisability of the results. Furthermore, the staggered design of the self-referral tool implementation at different time points accounted for any natural changes in referral numbers. Our pre- versus post-period design and comparison to matched NHS Talking Therapies services further ensured that the observed results are not due to general external effects or other time-dependent factors or confounding factors.

The findings of this research demonstrate the potential of AI-enabled conversational tools to facilitate the self-referral process for mental health services and make it more accessible for different groups of people. This is particularly timely in light of the World Health Organisation’s Global strategy on digital health [World Health Organization, 2021a] and many healthcare providers around the world setting goals, such as the NHS Long-Term Plan in the UK [National Health Service, 2019], aiming to increase the number of people accessing mental health services in accordance with the UN 2030 Sustainable Development Goals (SDGs). Our results further support the notion that digital technologies can support reaching the SDGs [Lancet, 2021] and support more people in accessing the services they need. Utilizing an AI-enabled self-referral tool could reduce barriers to access, resulting in more referrals and improved treatment efficacy [Rollwage et al., 2022] and a decrease in the burden on individuals and society.

5 Acknowledgments

We thank Emili Ivanova for her contribution to the figures and illustrations.

6 Competing interests

JH, SV, BC, RH and MR are employed by Limbic Limited and hold shares in the company. TH is working as a paid consultant for Limbic Limited.

7 Data and code availability

Code and data regarding the referral analysis will be made available upon submission acceptance. The qualitative feedback data will not be publicly available because the information could compromise participants’ privacy/consent.

References

- L. H. Andrade, J. Alonso, Z. Mneimneh, J. Wells, A. Al-Hamzawi, G. Borges, E. Bromet, R. Bruffaerts, G. De Girolamo, R. De Graaf, et al. Barriers to mental health treatment: results from the who world mental health surveys. *Psychological medicine*, 44(6):1303–1317, 2014.
- R. Bhakta, M. Savin-Baden, and G. Tombs. Sharing secrets with robots? In *EdMedia+ Innovate Learning*, pages 2295–2301. Association for the Advancement of Computing in Education (AACE), 2014.
- V. Braun and V. Clarke. Reflecting on reflexive thematic analysis. *Qualitative research in sport, exercise and health*, 11(4):589–597, 2019.
- J. S. Brown, J. Boardman, N. Whittinger, and M. Ashworth. Can a self-referral system help improve access to psychological treatments? *British Journal of General Practice*, 60(574):365–371, 2010.
- S. Clement, O. Schauman, T. Graham, F. Maggioni, S. Evans-Lacko, N. Bezborodovs, C. Morgan, N. Rüsch, J. S. Brown, and G. Thornicroft. What is the impact of mental health-related stigma on help-seeking? a systematic review of quantitative and qualitative studies. *Psychological medicine*, 45(1):11–27, 2015.
- P. W. Corrigan, B. G. Druss, and D. A. Perlick. The impact of mental illness stigma on seeking and participating in mental health care. *Psychological Science in the Public Interest*, 15(2):37–70, 2014.
- G. De Girolamo, J. Dagani, R. Purcell, A. Cocchi, and P. McGorry. Age of onset of mental disorders and use of mental health services: needs, opportunities and obstacles. *Epidemiology and psychiatric sciences*, 21(1):47–57, 2012.
- Department of Health and Social Care. Progress in improving mental health services in england, Feb 2023. URL <https://www.nao.org.uk/wp-content/uploads/2023/02/Progress-in-improving-mental-health-services-CS.pdf>.
- T. L. G. Health. Mental health matters. *The Lancet. Global Health*, 8(11):e1352, 2020.
- K. Jayaraajan, A. Sivananthan, A. Koomson, A. Ahmad, M. Haque, and M. Hussein. The use of digital solutions in alleviating the burden of iapt’s waiting times. *International Journal of Risk & Safety in Medicine*, (Preprint):S1–S8, 2022.
- J. A. Johnson, P. Sanghvi, and S. Mehrotra. Technology-based interventions to improve help-seeking for mental health concerns: a systematic review. *Indian Journal of Psychological Medicine*, 44(4): 332–340, 2022.
- M. Khanbhai, P. Anyadi, J. Symons, K. Flott, A. Darzi, and E. Mayer. Applying natural language processing and machine learning techniques to patient experience feedback: a systematic review. *BMJ Health & Care Informatics*, 28(1), 2021.
- L. Kola, B. A. Kohrt, C. Hanlon, J. A. Naslund, S. Sikander, M. Balaji, C. Benjet, E. Y. L. Cheung, J. Eaton, P. Gonsalves, et al. Covid-19 mental health impact and responses in low-income and middle-income countries: reimagining global mental health. *The Lancet Psychiatry*, 8(6):535–550, 2021.
- N. Koutsouleris, T. U. Hauser, V. Skvortsova, and M. De Choudhury. From promise to practice: towards the realisation of ai-informed mental health care. *The Lancet Digital Health*, 2022.
- K. Kroenke, R. L. Spitzer, and J. B. Williams. The phq-9: validity of a brief depression severity measure. *Journal of general internal medicine*, 16(9):606–613, 2001.
- Lancet. Can digital technologies improve health? *Lancet (London, England)*, 398(10312):1663, 2021.
- A. M. Loosen, V. Skvortsova, and T. U. Hauser. Obsessive-compulsive symptoms and information seeking during the covid-19 pandemic. *Translational psychiatry*, 11(1):1–10, 2021.
- E. Mahase. Workforce crisis has left mental health staff at “breaking point” as demand rises, 2020.

- M. Moitra, D. Santomauro, P. Y. Collins, T. Vos, H. Whiteford, S. Saxena, and A. J. Ferrari. The global gap in treatment coverage for major depressive disorder in 84 countries from 2000–2019: A systematic review and bayesian meta-regression analysis. *PLoS Medicine*, 19(2):e1003901, 2022.
- J. C. Mundt, I. M. Marks, M. K. Shear, and J. M. Greist. The work and social adjustment scale: a simple measure of impairment in functioning. *The British Journal of Psychiatry*, 180(5):461–464, 2002.
- National Health Service. Nhs long term plan, 2019. URL <https://www.longtermplan.nhs.uk/>.
- NHS Digital. Improving access to psychological therapies self-referral process research, May 2019. URL <https://digital.nhs.uk/publication-system/statistical/improving-access-to-psychological-therapies-key-performance-indicators-iapt-kpis/iapt-self-referral-process-research>.
- F. Ornell, W. V. Borelli, D. Benzano, J. B. Schuch, H. F. Moura, A. O. Sordi, F. H. P. Kessler, J. N. Scherer, and L. von Diemen. The next pandemic: impact of covid-19 in mental healthcare assistance in a nationwide epidemiological study. *The Lancet Regional Health-Americas*, 4:100061, 2021.
- A. Reichert and R. Jacobs. The impact of waiting time on patient outcomes: Evidence from early intervention in psychosis services in england. *Health economics*, 27(11):1772–1787, 2018.
- M. E. Rodríguez-Rivas, A. J. Cangas, L. A. Cariola, J. J. Varela, and S. Valdebenito. Innovative technology-based interventions to reduce stigma toward people with mental illness: Systematic review and meta-analysis. *JMIR Serious Games*, 10(2):e35099, 2022.
- M. Rollwage, K. Juchems, J. Habicht, B. Carrington, T. Hauser, and R. Harper. Conversational ai facilitates mental health assessments and is associated with improved recovery rates. *medRxiv*, 2022.
- S. Rose and M. J. Van der Laan. Why match? investigating matched case-control study designs with causal effect estimation. *The international journal of biostatistics*, 5(1), 2009.
- B. N. Rudd and R. S. Beidas. Digital mental health: the answer to the global mental health crisis? *JMIR Mental Health*, 7(6):e18472, 2020.
- D. F. Santomauro, A. M. M. Herrera, J. Shadid, P. Zheng, C. Ashbaugh, D. M. Pigott, C. Abbafati, C. Adolph, J. O. Amlag, A. Y. Aravkin, et al. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the covid-19 pandemic. *The Lancet*, 398(10312):1700–1712, 2021.
- R. L. Spitzer, K. Kroenke, J. B. Williams, and B. Löwe. A brief measure for assessing generalized anxiety disorder: the gad-7. *Archives of internal medicine*, 166(10):1092–1097, 2006.
- Z. Steel, C. Marnane, C. Iranpour, T. Chey, J. W. Jackson, V. Patel, and D. Silove. The global prevalence of common mental disorders: a systematic review and meta-analysis 1980–2013. *International journal of epidemiology*, 43(2):476–493, 2014.
- C. B. Taylor, E. E. Fitzsimmons-Craft, and A. K. Graham. Digital technology can revolutionize mental health services delivery: The covid-19 crisis as a catalyst for change. *International Journal of Eating Disorders*, 53(7):1155–1157, 2020.
- M. E. Te Pas, W. G. Rutten, R. A. Bouwman, and M. P. Buise. User experience of a chatbot questionnaire versus a regular computer questionnaire: prospective comparative study. *JMIR Medical Informatics*, 8(12):e21982, 2020.
- The National Collaborating Centre for Mental Health. The improving access to psychological therapies manual, August 2021. URL <https://www.england.nhs.uk/wp-content/uploads/2018/06/the-iapt-manual-v5.pdf>.
- J. Thome, J. Deloyer, A. N. Coogan, D. Bailey-Rodriguez, O. A. da Cruz E Silva, F. Faltraco, C. Grima, S. O. Gudjonsson, C. Hanon, M. Holly, et al. The impact of the early phase of the covid-19 pandemic on mental-health services in europe. *The World Journal of Biological Psychiatry*, 22(7):516–525, 2021.

- G. Thornicroft, C. Sunkel, A. A. Aliev, S. Baker, E. Brohan, R. El Chammay, K. Davies, M. Demissie, J. Duncan, W. Fekadu, et al. The lancet commission on ending stigma and discrimination in mental health. *The Lancet*, 400(10361):1438–1480, 2022.
- L. Tudor Car, D. A. Dhinakaran, B. M. Kyaw, T. Kowatsch, S. Joty, Y.-L. Theng, and R. Atun. Conversational agents in health care: scoping review and conceptual analysis. *Journal of medical Internet research*, 22(8):e17158, 2020.
- U. Volpe, A. Mihai, V. Jordanova, and N. Sartorius. The pathways to mental healthcare worldwide: a systematic review. *Current Opinion in Psychiatry*, 28(4):299–306, 2015.
- P. S. Wang, M. Angermeyer, G. Borges, R. Bruffaerts, W. T. Chiu, G. De Girolamo, J. Fayyad, O. Gureje, J. M. Haro, Y. Huang, et al. Delay and failure in treatment seeking after first onset of mental disorders in the world health organization’s world mental health survey initiative. *World psychiatry*, 6(3):177, 2007.
- M. E. Williams, J. Latta, and P. Conversano. Eliminating the wait for mental health services. *The journal of behavioral health services & research*, 35:107–114, 2008.
- L. Wilson and M. Marasoiu. The development and use of chatbots in public health: Scoping review. *JMIR human factors*, 9(4):e35882, 2022.
- World Health Organization. Global strategy on digital health 2020-2025. 2021a.
- World Health Organization. Who report highlights global shortfall in investment in mental health, 2021b. URL <https://www.who.int/news/item/08-10-2021-who-report-highlights-global-shortfall-in-investment-in-mental-health>.