

Exploring Patient-Generated Annotations to Digital Clinical Symptom Measures for Patient-Centered Communication

EUNKYUNG JO, University of California, Irvine, USA

RACHAEL ZEHRUNG, University of California, Irvine, USA

KATHERINE GENUARIO, University of California, Irvine, USA

ALEXANDRA PAPOUTSAKI, Pomona College, USA

DANIEL A. EPSTEIN, University of California, USA

Patients' self-reports are crucial for effective care management of clinical conditions involving subjective symptoms. While patients often value the ability to bring in different forms of self-report data to convey their lived experiences, they often struggle to make their data practically usable in clinical settings. To better center patient needs in communicating illness experiences in clinical contexts, we explore the idea of patient annotations to digital clinical self-report measures, specifically in the context of discontinuing antidepressants. Through interviews with 20 patients with AT Annotator, a digital aid to introduce the concept of annotations, we found that participants perceived annotations to digital clinical measures as a means to enrich self-report measures and reduce the cognitive and emotional burden of logging. However, concerns were raised regarding potential disruptions in patient-provider relationships and the sensitive and complex nature of mental health contexts. We discuss opportunities for annotations to promote patient-centered communication by balancing with clinical practicality and incorporating customization support for patients' communication needs.

CCS Concepts: • **Human-centered computing** → **Human computer interaction (HCI)**; **Empirical studies in collaborative and social computing**.

Additional Key Words and Phrases: Patient-Generated Health Data; Patient-centered communication; Personal Informatics; Self-Tracking; Annotations; Clinical self-report measures; Psychiatric Drugs; Antidepressants

ACM Reference Format:

Eunkyung Jo, Rachael Zehrung, Katherine Genuario, Alexandra Papoutsaki, and Daniel A. Epstein. 2024. Exploring Patient-Generated Annotations to Digital Clinical Symptom Measures for Patient-Centered Communication. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW2, Article 458 (November 2024), 26 pages. <https://doi.org/10.1145/3686997>

1 Introduction

Many clinical conditions involve subjective and idiosyncratic symptoms that affect individuals' daily lives in varying ways, including chronic pain [2, 3, 54, 83, 89], cancer [37, 50, 51, 62, 80], and mental illness [13, 43, 70, 71, 74]. In such cases, patients' self-reports—or any methods that rely on an individual's own descriptions of their symptoms, behaviors, feelings, and attitudes [39]—play a vital role in care management by allowing patients to communicate their illness experiences with their providers. Traditional methods to elicit patient self-reports include standardized paper-based

Authors' Contact Information: [Eunkyung Jo](mailto:eunkyu@uci.edu), University of California, Irvine, USA, eunkyu@uci.edu; [Rachael Zehrung](mailto:rzehrung@uci.edu), University of California, Irvine, USA, rzehrung@uci.edu; [Katherine Genuario](mailto:kgenuari@uci.edu), University of California, Irvine, USA, kgenuari@uci.edu; [Alexandra Papoutsaki](mailto:appa2017@pomona.edu), Pomona College, USA, appa2017@pomona.edu; [Daniel A. Epstein](mailto:epstein@ics.uci.edu), University of California, USA, epstein@ics.uci.edu.



This work is licensed under a [Creative Commons Attribution International 4.0 License](https://creativecommons.org/licenses/by/4.0/).

© 2024 Copyright held by the owner/author(s).

ACM 2573-0142/2024/11-ART458

<https://doi.org/10.1145/3686997>

measures, but relying on this approach risks losing valuable information about patients' lived experiences [50, 54] and hindering their agency in managing illness in everyday life [9, 10, 32].

To mitigate the limitations of the traditional approaches to eliciting patient self-reports, Human-Computer Interaction (HCI), Computer-Supported Collaborative Work (CSCW), and Health Informatics communities are increasingly developing and using technology to support patient self-reports. In particular, HCI and CSCW researchers have proposed novel methods such as pictorial trackers [9], digital storyboards [50, 51], and media probes [29, 50, 51] to account for patients' individual illness experiences. These studies showed that customizable tracking tools could lead to an increased sense of agency [9] and promote patient-centered communication [50, 51].

Although patients often value the ability to bring in these various formats for capturing data, they often face issues making their data practically usable in clinical settings, making it hard to communicate their lived experiences [27, 60, 104, 105]. While providers often use standardized self-report measures during intake [20] because they reliably capture aspects of patients' quality of life [35, 62], patients might feel their health experiences are not fully conveyed through them [50, 54, 62]. One way to better center patient needs in clinical symptom measures routinely used in clinical care is through **annotations**, which allow patients to adapt and extend these self-report measures with their own ideas. Allowing patients to annotate standardized self-report measures with free-form data could potentially empower them to convey their lived experiences to their providers while preserving clinical relevance.

In this paper, we explore the idea of patient annotations to clinical self-report measures as one way to promote patient-centered communication. We examine two research questions:

- **RQ1:** What goals do patients think annotations to clinical self-report measures support in promoting patient-centered communication?
- **RQ2:** What goals do patients think annotations to clinical self-report measures may fall short of achieving in promoting patient-centered communication?

We examine these research questions in the context of discontinuing antidepressants. More than half of the people who attempt to discontinue antidepressants experience debilitating withdrawal symptoms [34], including flu-like symptoms, insomnia, nausea, and sensory disturbances [34, 85, 86, 103]. Since most withdrawal symptoms lack objective measures, providers typically rely on patients' self-reports using standardized tools. However, due to the subjective nature of the withdrawal symptoms and variability of patient tolerance for dose reductions [52, 86], such measures might not fulfill providers' and patients' symptom monitoring needs. Discontinuation of antidepressants thus presents a useful case study for studying patient adaptations of clinical self-report measures.

To understand how technology can better support patient adaptation to clinical measures, we designed AT Annotator, a low-fidelity prototype that allows users to enhance standardized clinical self-report measures with five annotation types: free-text notes, emojis, animated GIFs, icons, and body parts. We then conducted individual interviews with 20 patients who were either interested in stopping their antidepressants or in the process of discontinuing them, using AT Annotator as a digital aid to introduce the concept of annotations. Participants felt that annotations could effectively support many of their communication goals by enriching clinical measures with their individual illness experiences, such as symptom fluctuations and the impact of medication changes on daily life. They believed annotations could also empower them to highlight their primary concerns to their providers, aid in recalling symptom experiences, and alleviate the logging burden. However, participants were concerned that visual annotations, such as GIFs and emojis, might disrupt the professional relationship with their providers due to their casual or ambiguous nature. Participants further pointed out that annotations might not adequately address the sensitivity

and complexity of mental health contexts, potentially imposing significant mental and emotional burdens while lacking the flexibility needed to convey mental health experiences.

Based on the findings, we discuss opportunities for annotations to support patients in conveying their personal experiences and influencing the direction of their communication with providers. We further propose incorporating customization support to cater to patients' diverse communication needs around the form of patient-generated data. Lastly, we highlight opportunities to enhance the clinical practicality of digital symptom measure annotations by guiding patients to focus on those more pertinent to their ongoing care when sharing them with providers.

The key contributions of this work include:

- Understanding of patient perceptions of what goals annotations to clinical symptom measures support or fall short of achieving in supporting patient-centered communication. While patients felt annotations would enrich clinical measures with their individual symptom experiences and alleviate the cognitive and emotional burden of logging, they also thought that visual annotations, such as GIFs and emojis, might interrupt the professional relationship with providers and fall short of considering the sensitivity and complexity of mental health contexts.
- Implications for further research and implementation of self-tracking technology for patient-centered communication, particularly around (1) examining the utility and challenges of annotations in real-world settings to understand the practical considerations, acceptance, and their role in patient empowerment, (2) incorporating customization support to cater to patients' diverse communication needs around the form of patient-generated data, and (3) balancing patient-centric communication with clinical practicality by offering guidance about what types of information could be more pertinent to the patient's ongoing care.

2 Background on Antidepressant Discontinuation

With the rise in diagnoses of mental health disorders in the United States, the prescription of psychiatric drugs has rapidly increased [1]. One in six adults in the United States receives a prescription for one or more psychiatric drugs per year, with antidepressants being the most commonly prescribed class of psychiatric medication [4, 81]. Recently, clinical guidelines have recommended discontinuing antidepressants when patients achieve complete symptom remission over an extended period of time. According to the American Psychiatric Association, antidepressants can be stopped in stable patients, although the precise timing has not been specified [8]. The United Kingdom's National Institute for Health and Care Excellence suggests stopping an antidepressant six months after symptom remission is achieved [41]. Discontinuation of antidepressants is also considered in other circumstances, such as when patients report intolerable side effects [86], when the medication is ineffective in treating a target condition [25], or when special conditions (e.g., pregnancy or breastfeeding) exist that may adversely affect ongoing antidepressant treatment [103].

Research has indicated that discontinuing antidepressants involves significant challenges. A systematic review article reported that more than half of the people who attempt to discontinue antidepressants experience withdrawal symptoms [34]. Patients may experience debilitating withdrawal symptoms, including flu-like symptoms, insomnia, nausea, and sensory disturbances [34, 85, 86, 103]. These symptoms can last for months or even years [34, 85], which could lead to serious psychiatric problems such as suicidal ideation [46]. Clinical recommendations advise gradual discontinuation (a taper) rather than abrupt cessation to prevent withdrawal symptoms [8, 41, 75]. However, there is no established protocol for planning the gradual discontinuation of antidepressants [52, 76]. The American Psychiatric Association [8] and the National Institute for Health and Care Excellence [41] only provide vague guidance on tapering, such as "*over the course of at least several weeks*" or "*at a*

rate proportional to the duration of treatment,” respectively. Therefore, providers devise tapering protocols based on their clinical intuition and often struggle if lacking experience [56].

Past work suggests that tapering plans should be tailored to individual patients along with careful symptom monitoring because tolerance for dose reductions could vary by individual [42, 45, 52, 86]. Providers typically rely on patients’ self-reports using clinician-rated instruments, such as the Discontinuation-Emergent Signs and Symptoms (DESS) checklist [88], to monitor whether patients have any new or worsening withdrawal symptoms or the Hamilton Depression Rating Scale (HDRS) [47], the Patient Health Questionnaire (PHQ-9) [61] and the Beck Depression Inventory (BDI) [14] to monitor potential relapse of depressive symptoms. If relapse occurs, providers often revert to a higher dose or consider alternative treatments such as different medications.

However, given the subjective nature of common withdrawal symptoms as well as depressive symptoms, this approach may be limited in capturing patients’ lived experiences during the taper. Recent work shows that many patients feel that they do not receive adequate support for their safe discontinuation of psychiatric drugs from their providers and turn to online health communities [42, 45, 79]. These patients often engage in self-tracking of their drug history and daily symptoms, with some following taper plans they found in online health communities. Acknowledging the importance of understanding patients’ lived experiences and the need for provider engagement in the tapering process [45], Papoutsaki et al. argued that such patient-generated logs could instead be leveraged to facilitate patient-provider collaboration for the safe discontinuation of psychiatric drugs [79]. For example, patient-generated logs could serve as evidence of patients’ illness experiences and medication adherence, allowing providers to distinguish between withdrawal and relapse and adjust their prescriptions accordingly.

3 Related Work

3.1 Patient-Generated Data and Self-Reports

Research in HCI, CSCW, and Health Informatics has explored ways to promote patient-centered communication using patient-generated health data, highlighting the benefits of patient-driven practices. Patient-tracked data can empower patients to assert their voice in clinical consultations by helping them construct and share their narratives [28, 82]. For example, it can serve as objective evidence of their illness experiences, which can be more acceptable for providers than a verbal narrative [82]. Prior work also showed that patient-tracked data could help providers better understand patients by providing rich information about patient preferences, goals, and values [27, 28, 53].

An important part of patient-generated health data is patient self-reports—including pain self-assessment [2, 3, 54, 83, 89], mood tracking [13, 50, 51, 70, 71, 74], and journals about daily routines [50, 51, 62]. Patients’ self-reports play a vital role in care management, enabling them to share qualitative insights about their illness experiences and improving providers’ understanding of their quality of life, well-being, and burden of symptoms [51, 105]. This information is particularly useful for understanding subjective and unique symptoms that impact daily life in conditions such as chronic pain [2, 3, 54, 83, 89], cancer [37, 50, 51, 62, 80], and mental illness [13, 43, 70, 71, 74].

Traditionally, self-reports are generated through clinical interviews conducted by trained professionals based on disease-specific measures [27, 78, 83, 98]. Although clinical interviews can ensure reliable patient assessment [66, 78], this approach has several drawbacks. These interviews rely on patients’ retrospective recall, which might be biased by their symptom experience at the time of reporting or recent extremes [80]. In addition, time constraints of typical clinical visits may prevent providers from eliciting detailed accounts of patient experiences [66, 78, 83]. Thus, providers often find clinical interviews insufficient for understanding what happens between visits [51].

Another common method to elicit patient self-reports is using self-report measures. Two terms closely related to self-reporting in clinical settings are patient-reported outcome (PRO) measures and observations of daily living (ODLs). PRO measures are validated tools for monitoring and assessing patients' symptoms, functional status, and health-related quality of life [20]. Numerous PRO measures exist, including disease-specific and generic ones, typically in the form of multiple-choice questionnaires [20]. Typical PRO measure questions include the frequency and severity of symptoms, their impact on daily life, and perceptions of conditions or treatments [35]. PROs are the gold standard for assessing symptoms in clinical trials [24, 92] and are also widely used in routine clinical care to monitor symptoms and evaluate treatment outcomes [35, 92]. As standardized, validated instruments, PROs reliably capture some aspects of patients' quality of life [35, 62]. However, relying on standard measures risks losing valuable nuances about patients' lived experiences [50, 54, 62] and diminishing their sense of agency in managing illness in everyday life [10, 11, 32]. Pichon et al. [82] suggested that patients could feel unheard or dismissed when standardized clinical measures used in their care fail to capture key aspects of their illness experiences.

In contrast to PROs, which are clinician-defined and patient-generated, ODLs are patient-defined and patient-generated data [22] with a greater focus on personal experiences of health. Brennan et al. highlighted that medical language often differs from patients' everyday experiences, suggesting ODLs as a new mechanism for patients to collect personally meaningful health data and express their unique experiences in their own words [22]. However, such patient-defined tracking often produces free-form data that may not align with clinical standards and are often seen as clinically irrelevant to providers [104, 105]. The heterogeneity in data forms and representations further complicates provider interpretation of patient-tracked data [27, 60, 104, 105]. In this work, we explore how tracking technology might better center patient needs in communicating their illness experiences with their providers while maintaining clinical relevance.

3.2 Patient-Centered Communication and Care

Research has suggested that patients and providers tend to focus on different aspects of illness, leading to divergent concerns and priorities [5, 7, 26, 48, 64, 84]. While providers tend to base their care recommendations on biomedical data (e.g., lab results), patients experience illness as personal events that directly impact various aspects of their daily lives [5–7, 26, 48, 64]. For example, patients may prioritize certain activities (e.g., gardening) or relationships (e.g., newborn grandchildren) when planning treatments, but these personal needs are not routinely discussed in clinical settings [64, 65]. Patients often struggle to communicate personal aspects of illness, withholding what matters most to them (i.e., values) due to a perceived communication boundary, which limits providers' ability to align care recommendations with patients' priorities [65]. In addition, providers may overlook non-medical aspects of illness that are important to patients due to their focus on medically oriented goals [15, 69]. Such communication gaps can lead to conflicts in care priorities, negatively influencing patient health outcomes [17].

Prior work suggests that aligning patient and provider perspectives requires a collaborative process to identify and translate meaningful patient concerns into actionable insights for providers [5–7]. Studies have proposed ways technology could help bridge the gap between patient and provider perspectives, particularly by empowering patients to communicate the unique, personal aspects of their illness experiences. Berry et al. [16, 17, 19] suggested utilizing open-ended prompts [16, 18] or visual artifacts (e.g., photos and videos) [17, 19] to better equip patients to discuss their values during clinical visits. They argued that open-ended, exploratory tools could help overcome perceived communication boundaries, whereas tools explicitly designed to generate discussion topics with their providers might reinforce existing boundaries, highlighting the value of supporting exploratory reflection [18]. Chung et al. [28, 29] characterized personal tracking data in

patient-provider collaboration as boundary negotiating artifacts [63] to mediate different spheres of expertise. Consistent with Berry et al. [15, 17, 19], they suggested the potential benefits of incorporating visual artifacts to resolve disagreements between patients and providers [28, 29].

Prior work has also proposed different types of technology features that support the curation of patient-driven tracking data, making it easier for providers to review meaningful data points within the time constraints of clinical visits. Filtering features can support patients and providers in focusing on tracking data that are most relevant to their health goals [29, 105]. Features for flagging or sorting based on a range of factors, such as the level of patient concerns or severity of symptoms, can also aid in communicating patients' priorities during visits [6, 7, 51, 82, 91]. Supporting annotations, or allowing patients to augment standardized self-report measures with their own ideas, could further enable patients to mark exceptional events and add contextual information to set an agenda before visits [6, 7, 28, 93]. However, we have a limited understanding of how patients and providers perceive these adaptations to better convey their illness experiences. As a first step to address this gap, we explore patient annotations to clinical measures as a means of facilitating patient-centered communication on subjective illness experiences.

3.3 Patient-Driven Tracking

Prior work has demonstrated that patients often track aspects of their illness experiences with or without provider involvement. Patients' data practices can provide useful self-insights into their conditions [40, 44, 74] and facilitate self-expression and mindfulness [3, 74]. Studies show that patients may prefer configuring their own tracking regimens rather than relying on pre-defined options from providers [9, 37, 50, 51, 72, 91]. Past work has proposed ways to support customized tracking for patient needs. For example, Schroeder et al. [94] proposed goal-directed tracking, allowing patients to customize data types and frequency based on their health goals. Similarly, Karkar et al. [57] presented TummyTrials, a tool that generates customized study protocols to identify causal relationships between food triggers and symptoms of irritable bowel syndrome.

Further, HCI studies have proposed various ways to support free-form descriptions of patients' illness experiences. Yamashita et al. [106] and Luo et al. [68] showed that free-form texts enable individuals to describe their experiences in a flexible manner. Stawarz et al. [99] and Mishra et al. [73] and other studies in medical informatics [97, 100] demonstrated that emojis could be an effective way to convey emotions in clinical settings as it reduces cognitive burden by using symbols frequently used in everyday life [97]. Icons have been suggested as a way to allow for succinct and efficient communication on teletherapy platforms [87]. Adams et al. [3] proposed tailoring pain assessment measures (e.g., coloring, number picker, illustrative faces, slide bar) to consider patients' individuality in how they experience and process symptoms. Ayobi et al. [9] suggested pictorial symptom journaling to improve a sense of agency in patients with multiple sclerosis, surfacing the need to consider individual data collection preferences. Hong et al. [50, 51] similarly proposed digital storyboards for adolescent cancer patients to communicate their illness experiences in their preferred ways. Prior work suggests that self-tracking can serve as a mechanism for patients to articulate their needs in personalized ways. Inspired by prior work, we explore how technology can better support patient-driven tracking through annotations to clinical measures.

4 Methods

To understand how annotations could support patient adaptation to clinical measures, we conducted interviews with 20 patients who were either interested in stopping or in the process of discontinuing their antidepressants. This study was approved by our institution's Institutional Review Board.

4.1 Study Procedures

We designed AT Annotator as a digital aid for understanding how annotations to clinical measures might help or hinder patients' communication needs when tapering antidepressants. The aim of designing AT Annotator as a low-fidelity prototype was to introduce the concept of annotations during the interviews.

4.1.1 Prototype Design. We designed a low-fidelity prototype, AT Annotator, to elicit conversations about how annotations to clinical measures might support their communication goals with their providers. The design of our prototype was influenced by findings from prior studies on clinicians' use of validated scales, both in the space of tapering antidepressants [14, 47, 61, 88] and outside [30, 31]. These studies suggested that providers regularly ask patients to complete clinical self-reports, review them prior to patient interactions, and use them to inform treatment plans. While AT Annotator was created for the specific case of antidepressant tapering, we expect the design principles could be applied to other domains where standardized clinical measures are used, such as chronic pain [30] or cancer [31].

Our prototype illustrated the following features:

- (1) *Standardized clinical self-report measures.* AT Annotator includes self-report questions that are widely used in the context of tapering antidepressants as well as part of standardized clinical measures, such as the PHQ-9 [61] and the DESS checklist [88] (Figure 1). PHQ-9 [61] is a widely used self-report measure designed to assess the severity of depressive symptoms. This scale is relevant to tapering antidepressants because it is useful for monitoring the relapse of depressive symptoms with dose changes. The DESS checklist [88] is a self-report measure specifically designed to monitor withdrawal symptoms that patients might experience when discontinuing antidepressants. This scale is useful for monitoring what type of withdrawal symptoms, if any, patients are experiencing and how severe they are. We included these questionnaires as prior work suggested that providers often monitor withdrawal symptoms during the antidepressant tapering process [56]. To ensure their relevance, a board-certified psychiatrist validated these questions as representative of the most frequent and noteworthy symptoms in the context of tapering antidepressants. In addition, we added some questions as a daily questionnaire, including daily mental wellbeing, the severity of common withdrawal symptoms (e.g., dizziness), the time when the symptoms were experienced, and medication adherence, with the advice of the board-certified psychiatrist.
- (2) *Annotations to clinical self-report measures.* AT Annotator provides users with the ability to enhance their clinical survey responses using five annotation types, which include free-text notes, emojis, animated GIFs, icons, and body parts (Figure 2). Prior studies in HCI and Health Informatics have proposed that these forms could effectively support aspects of self-tracking and patient-provider communication. For example, free-form texts provide the flexibility for users to describe their thoughts and feelings [68, 106] and convey contextual information around standardized clinical measures [95]. Emojis and animated GIFs have been suggested as a way to better convey emotions within and outside clinical settings [12, 33, 55, 73, 97, 99, 100]. Similarly, icons have been suggested as a way to provide communication shortcuts on teletherapy platforms [87]. With free-text notes, users can type out anything that they want to convey to their doctor in free form. With emojis, GIFs, and icons, users can search for relevant images and add them to describe how they feel about particular symptoms or their overall health. With body parts, users can specify the affected area by clicking on a specific part of the body manikins (Figure 2). Users can add these annotations to their clinical survey responses by clicking on the edit icon next to each question (Figure 1 (b)) and selecting an annotation method in the bottom navigation bar (Figure 1 (c)). Given our study consisted

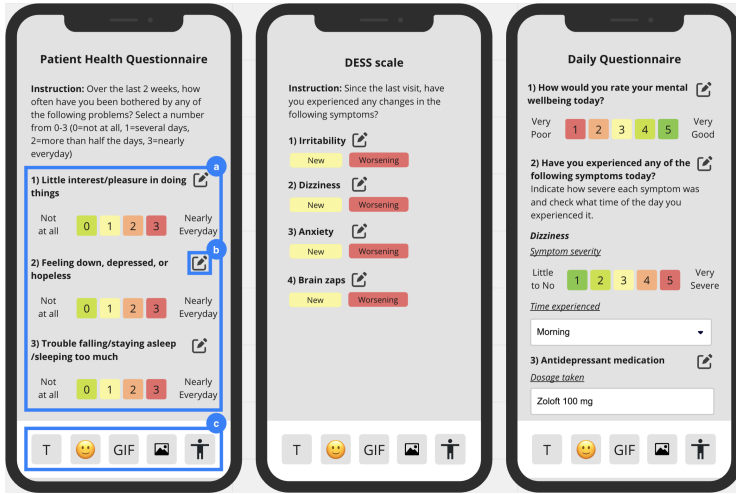


Fig. 1. The AT Annotator low-fidelity prototype includes (a) Standardized self-report questions including the PHQ-9 [61] and the DESS checklist [88], (b) Features for adding annotations to responses for specific questions by clicking on the edit icon next to each question, (c) Different annotation methods in the bottom navigation bar, including free-text notes, emojis, animated GIFs, icons, and body parts.

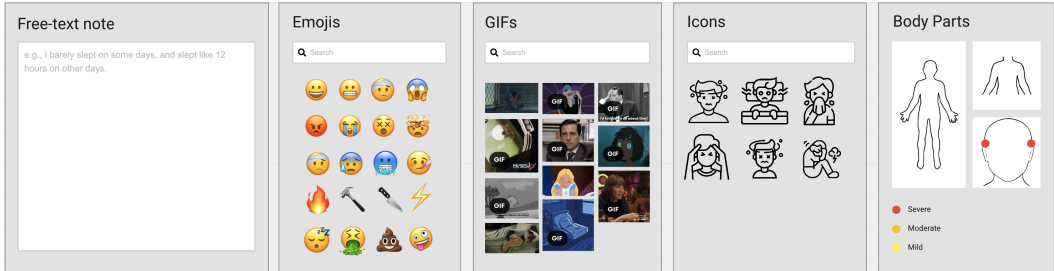


Fig. 2. The AT Annotator low-fidelity prototype contains five types of annotation methods: free-text notes, emojis, animated GIFs, icons, and body parts. Users could reflect on the utility of adding these annotations to each question of their clinical survey responses. Icons and GIFs are from [GIPHY.com](#) and [Flaticon.com](#).

of exploratory interviews with a low-fidelity prototype, the visual artifacts (e.g., emojis, GIFs, icons) in Figure 2 were mere examples that show the overall concept of annotations to clinical measures, rather than specific design outcomes to be evaluated. Our goal for their representation was to aid participants in envisioning the use of annotation tools.

- (3) *Symptom summary report*. AT Annotator generates a monthly symptom report to provide a summary of the annotations as well as their responses to the clinical measures for facilitating patient-centered communication about their illness experiences during or between visits. These sorts of summary reports have been commonly proposed and created to present patient-generated data to clinicians [29, 59, 82, 93, 104, 105]. Some common features of these summary reports include numerical summaries and graphs to allow providers to get a quick overview of the data and decide whether they should take a close look at it if they find anything noteworthy, which were replicated in AT Annotator. Figure 3 shows an example of a summary of annotations for an entire month. As prior work suggested that providers often decrease the medication dosage every month [56], we chose a month-long period to summarize patient-generated data through AT Annotator. The report provides (1) users'

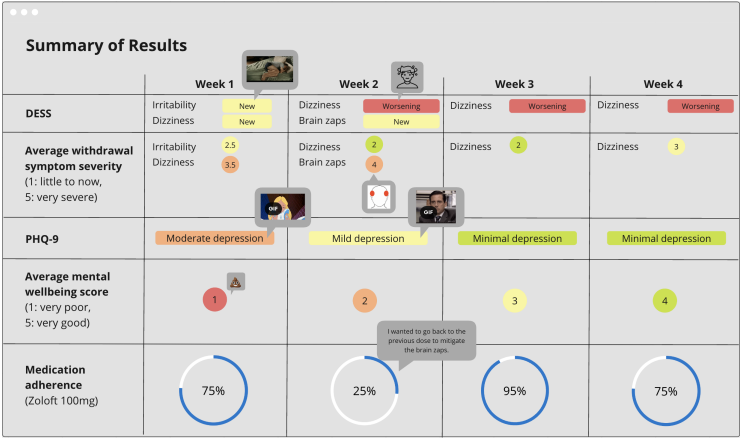


Fig. 3. The AT Annotator prototype included a monthly symptom report that summarizes the annotations as well as responses to the clinical measures. The report provides (1) user responses to standardized questionnaires, such as DESS [88] and PHQ-9 [61], (2) average withdrawal symptom severity and mental well-being scores, (3) medication adherence, and (4) annotations added by users. The report utilizes color-coded indicators; green denotes positive trends in symptoms, while red indicates negative trends. Annotations appear in gray callouts attached to corresponding clinical survey responses. Icons and GIFs are from [GIPHY.com](https://www.giphy.com) and [Flaticon.com](https://www.flaticon.com).

responses to standardized questionnaires, such as DESS [88] and PHQ-9 [61], (2) average withdrawal symptom severity and mental well-being scores, (3) medication adherence, and (4) annotations added by users. The report utilizes color-coded indicators; green denotes positive trends in symptoms, while red indicates negative trends. Annotations appear in gray callouts attached to corresponding clinical survey responses.

4.1.2 Interview Study. We conducted all interviews remotely using video conferencing. One or two members of our research team conducted each interview, lasting 50 to 60 minutes. Our study used AT Annotator as a backdrop to understand the role of technology in communicating patient experiences to providers during the tapering process.

During the interviews, we first asked about participants’ experiences of stopping their antidepressants, focusing on existing interactions with their providers and challenges with changes to their medication doses. Next, we showed an example of standardized clinical measures commonly used in the context of tapering antidepressants in the low-fidelity prototype, asking what participants felt about the topics covered by the survey and if there were other things that they may want to convey to their doctors. We then walked participants through different annotation methods (i.e., free-text notes, emojis, GIFs, icons, and body parts) in the prototype, asking whether and how different annotation methods might help patients convey their individual experiences to their providers. We emphasized to participants that all described annotations (e.g., the specific emojis and GIFs) shown in the prototype were mere examples of the overall types of annotations and encouraged them to think broadly about how they might perceive the utility of that type of annotations. Lastly, we showed participants a prototype of a monthly report that provides a summary of the patient annotations as well as their responses to the clinical measures, asking their opinions about how such a report might impact communication with their providers.

4.1.3 Data Analysis. All interviews were video-recorded, automatically transcribed, and manually revised to correct errors afterward. We used inductive thematic analysis [21, 90] to qualitatively

analyze interview transcripts without preconceived categories or theories. The first author open-coded the transcripts to identify patterns in the dataset. The full research team discussed and identified themes through multiple rounds of peer debriefing meetings. From this coding, we identified the main themes about patient goals being supported or left unmet by annotations, which we used to organize our findings. The final codebook contained six parent codes—including concerns about tapering, limitations of existing clinical communication practices, enriching clinical symptom measures, alleviating logging burden, establishing a professional communication environment, and considering the sensitivity and complexity of mental health contexts—and 20 child codes.

4.2 Participants

We recruited participants through three channels: (1) ResearchMatch,¹ a non-profit program that helps connect people interested in clinical studies and researchers (17 participants), (2) recommendations of psychiatric providers at the medical center of our institution (two participants), and (3) word-of-mouth through our personal networks (one participant). Our eligibility criteria included individuals who are: (1) 18–65 years old, (2) have been receiving treatment with an SSRI, SNRI, or other SRI antidepressants for at least three months, (3) are either interested in stopping their antidepressants or in the process of discontinuing them, and (4) do not have a comorbid DSM-5 diagnosis of schizophrenia, schizoaffective disorder, or a substance use disorder that is currently active. We settled on these criteria with the advice of the board-certified psychiatrist in order to (1) include patients who could consent and had relevant interests and experiences, (2) exclude patients who had a different serious mental health condition (e.g., schizophrenia) and were receiving psychiatric treatment that would benefit from monitoring additional clinical self-reports, as this could distract from our RQs, and (3) exclude patients who would require more active provider involvement than the typical, which would reduce the need for at-home clinical monitoring. We did not exclude patients who had other multiple chronic conditions. We compensated each participant with \$30 cash or a gift card for a one-hour interview session.

Participants included 18 Caucasians, one Asian, and one Hispanic or Latino, consisting of four males and 16 females, ranging in age from 22 to 65 (median = 35.5) (Table 1). The median annual household income of our participants fell between \$50,000 and \$75,000. The majority of participants had been taking antidepressants for over four years (N=12). Four participants were in the process of discontinuing their antidepressants at the time of the study, twelve participants had prior experiences of attempting to discontinue their antidepressants, and four participants had no prior experience of discontinuing their antidepressants but were interested in doing so. Most participants (12) had taken antidepressants for over four years, while one participant had taken them for 3–4 years, five for 2–3 years, and two for 1–2 years. All our participants were based in the United States.

4.3 Limitations

Our study sample skews toward the experiences of Caucasian females in their 20s and 30s, potentially influencing the transferability of our findings [36]. According to the National Center for Health Statistics [23], the prevalence of antidepressant use was significantly higher among females (17.7%) than males (8.4%) during 2015–2018, indicating a nearly twofold difference. Thus, our study sample somewhat reflects the gender distribution of the overall population taking antidepressants. Males, older individuals, and people from diverse racial backgrounds may have distinct relationships with their healthcare providers and different perspectives on annotation methods than those identified in our study. For example, males may experience unique challenges in healthcare settings due to social expectations of masculinity, affecting their communication style and openness with healthcare

¹<https://www.researchmatch.org/>

Table 1. Participant demographics, including gender, age, years on antidepressants, and tapering attempts

Alias	Years on antidepressants	Tapering attempts
P1 (F, 65)	≥ 4	Ongoing
P2 (M, 36)	3–4	3–4 years ago
P3 (F, 32)	≥ 4	In the last year
P4 (M, 57)	≥ 4	None
P5 (M, 25)	≥ 4	Ongoing
P6 (F, 53)	≥ 4	2–3 years ago
P7 (F, 22)	≥ 4	2–3 years ago
P8 (F, 23)	2–3	In the last year
P9 (M, 65)	≥ 4	None
P10 (F, 23)	2–3	Ongoing
P11 (F, 22)	≥ 4	More than 4 years ago
P12 (F, 26)	2–3	Ongoing
P13 (F, 38)	≥ 4	In the last year
P14 (F, 38)	≥ 4	2–3 years ago
P15 (F, 35)	≥ 4	More than 4 years ago
P16 (F, 48)	≥ 4	More than 4 years ago
P17 (F, 51)	1–2	None
P18 (F, 38)	2–3	In the last year
P19 (F, 29)	1–2	None
P20 (F, 27)	2–3	More than 4 years ago

providers [101]. Older individuals are more likely to have established long-term relationships with their healthcare providers and be more accustomed to traditional healthcare practices. Therefore, they might be less receptive to annotation methods that heavily rely on technology or less interested in adopting a novel tool that has not yet been integrated into traditional clinical practices.

In order to understand the utility of adapting provider-administered clinical scales, we narrowed our focus to the perspectives of patients and examined if such an approach could meet their needs. While this initial exploration is valuable, engaging with healthcare providers is likely to reveal potentially conflicting viewpoints. Given providers' scarce resources [27, 104, 105], we expect there are serious challenges to adopting this approach in clinical settings despite the opportunity for improved patient care. Future research on incorporating provider needs into the design of annotation tools would advance our understanding of how to present the annotations in ways providers could understand and accept, and whether providers would engage with them at all.

Our study method was exploratory interviews with the aid of a low-fidelity prototype, AT Annotator, to elicit patient perspectives on the potential use of annotation tools in their care processes. Therefore, the findings were grounded in participants' envisioned use of tools that incorporate such features instead of their real-world experiences. Although our low-fidelity prototype was useful for exploring how people envision using annotations for patient-centered communication, deploying a functional prototype in real-world settings would be necessary to examine how often patients would annotate, what sorts of annotations they would include, and how they would use the annotations to communicate with their providers in practice. Further, before deploying such a tool in real-world clinical settings, all contents and visual artifacts should be carefully designed to address specific patient needs and validated by providers who have domain expertise.

5 Findings

We present our findings on how annotations might support patient-centered communication in the context of tapering antidepressants, focusing on whether patients believed annotations to clinical self-reported measures could support their communication goals in clinical settings.

5.1 Goals Supported by Annotations

Overall, participants perceived that annotations to clinical measures could effectively support many of their goals in communicating with their providers, particularly by enriching clinical measures with individual symptom experiences and alleviating the cognitive and emotional burden of logging.

5.1.1 Enriching Clinical Measures with Individual Symptom Experiences. Participants perceived that annotations could enrich clinical measures with individual symptom experiences, including (1) symptoms not covered in standard surveys, (2) localized physical symptoms, (3) symptom severity, frequency, and duration, (4) personal circumstances impacting mental health, (5) complex mental health symptoms, and (6) the impact of medication changes on daily life.

Conveying symptoms that are not covered in standard clinical surveys. Several participants mentioned experiencing symptoms not covered in commonly used standardized clinical self-report measures during antidepressant tapering. For example, P2 had sexual side effects when taking antidepressants and desired to track his improvement as the medication dose decreased, but the clinical surveys he was asked to fill out during appointments lacked relevant questions. Participants also highlighted some aspects of psychiatric symptoms not covered in the clinical surveys. P13 was concerned about communicating her itching, which she had experienced during her previous episodes of medication nonadherence: “My doctor says that itching can be a physical manifestation of depression. If I don’t take it for a day or two by accident, I start getting itching and it’s just horrible. I wish the survey covered those physical aspects.” P15 similarly wished that there was a way for her to communicate night terrors with her provider: “There are things that are not part of the survey that I would talk to my doctor about. Sometimes I experience night terrors, and there’s never any kind of question about that.” Therefore, she appreciated annotations, specifically free text, as a means to communicate symptoms not covered by clinical measures.

Describing localized physical symptoms. Participants thought annotations with body parts would help them describe localized physical symptoms. P2 believed that visual annotations that indicate specific body parts would help him convey withdrawal symptoms more accurately: “Sometimes you’re having an issue in a certain part of the body, and you have a hard time explaining where an issue is on the body. For example, when I started tapering off my medication this time, I felt some pressure on my head. It would have been helpful to be able to say that these are the locations where I’m feeling this pressure and be able to annotate that with the graphic.” P20 further emphasized the value of visual annotations, particularly in telemedicine settings: “You could annotate where the headache is and pinpoint exactly where that’s happening, especially for telemedicine since you’re not physically with your doctor to tell them where things are happening.”

Conveying symptom severity, frequency, and duration. Participants found clinical measures inadequate in accurately conveying details about the severity, frequency, and duration of symptoms during the tapering process, viewing annotations as an effective means to describe such details. P20 expressed concerns about solely relying on clinical measures to summarize her psychiatric symptoms between visits because it would not accurately reflect fluctuations in symptoms: “I don’t want them to just get a view of the week and be like, okay, she was three out of five on average. It’s not a fair assessment for that week. I might have had three really great days and four terrible days. The only

thing that they [providers] are seeing would be the average. That gives me anxiety.” Participants thus perceived the benefits of using free-text annotations in accurately conveying details about their psychiatric symptoms: *“If I felt like the survey options didn’t fully encapsulate what my experience was, I would use the notes to clarify. Let’s say I select, ‘I felt depressed nearly every day,’ but in the free text note, I could specify, ‘Yes, I felt it every day, but it didn’t last very long. Most days, it lasted a short period of time.’ So I could use that to give some more context to my answer. (P12)”*

Describing personal circumstances impacting mental health. Participants further desired to convey personal circumstances that may contribute to their deteriorating mental health during the taper. However, standard clinical surveys did not allow them to communicate recent major life events that likely impacted their mental health, suggesting the potential benefits of annotations in facilitating such communication. P1 said: *“One thing [that I would add] is any life changes. Found out my son is an addict, got divorced, mother died. Any life changes can play into this.”* They thus envisioned that annotations could help convey significant life events as relevant contexts to their clinical survey responses: *“I think it [annotations] adds a lot more nuances to the results. I don’t think that I could get my entire message across without being able to say like, yes, I had moderate depression during week one. But my cat just died. That’s what brought down my mood. Maybe not the medication. (P11)”*

Participants further emphasized the need to convey daily events that may have influenced their mental health during the taper. P18 wished to convey personal circumstances surrounding her mental health flare-ups: *“Contexts of what was happening around me when the symptoms were coming up is the part that I would be trying to get across most. If both of my kids were sick and it was a particularly stressful day, I want to contextualize the symptoms more.”* P14 expressed concerns about reporting her current mental health status solely through standard clinical surveys because it could misguide her provider in making care decisions: *“If I hit three [nearly every day] for little interest/pleasure in doing things, they [providers] will only see three. But there might have been some things behind it, like suffering from a UTI [urinary tract infection]. If I cannot explain why I picked three, I worry I’ll be seen as my meds aren’t working, I’m a horrible human being, and nothing helps. For me, not being able to explain why I chose three increases anxiety.”* Participants thus highlighted that annotations to clinical surveys could help them convey potential factors influencing their mental health, aiding providers in making informed decisions about medication changes.

Articulating complex mental health symptoms. Participants perceived that annotations could help them articulate complex mental health symptoms during the taper. They often mentioned experiencing challenges in explaining complex psychiatric symptoms to their providers, wishing to have alternative ways to express their mental health status. P4 said: *“There are always complications in trying to explain to someone how you feel mentally, especially if that’s someone you rely on for medical advice.”* P20 similarly mentioned challenges in describing her mental health status: *“I guess the hardest part is to describe how you’re feeling. When I think of how I’m feeling about the medication changes, ‘good’ doesn’t necessarily cover everything. It’s really difficult to describe.”*

Some participants believed free-text annotations would add nuance to clinical survey responses. P12 said: *“When you’re just looking at the survey answers without any of those annotations, it feels kind of robotic, like you’re trying to squeeze your feelings into boxes. I think free-text notes could provide nuances during the [tapering] process.”* Other participants perceived that visual annotations would better convey complex mental health states: *“I could see myself using GIFs to portray some feelings, just because it is sometimes easier to have visuals rather than put it into words. Using a GIF could be helpful when you don’t have the words for certain things. (P20)”* Participants felt that conveying personal aspects of their mental health through annotations would help their providers understand them better. P14 noted: *“We want someone to understand how we’re feeling. Emojis or GIFs could help my doctor be like, okay, that’s how she’s feeling right now. I get that. I physically see it.”*

P19 added that annotations might lead her provider to ask more personalized questions: *“It would be more specific to my personal needs than just the general questions that they typically ask. So my doctor will know what would be a better question when they see these versus when I talk to them.”*

Conveying the impact of medication changes on daily life. Participants perceived annotations as a valuable tool for describing the impact of medication changes on their daily lives, assisting providers in making informed tapering decisions. P3 thought tolerability of her withdrawal symptoms would be important information for her provider: *“Is your anxiety tolerable with decreasing your dose? If you’re dizzy, is it tolerable dizzy? Can you still function? Can you get yourself to the bathroom? Kind of wondering how my doctor would change my dose or plan based on this scale.”* Since this information was not covered by clinical measures, P3 wanted to use icons for communicating the tolerability of withdrawal symptoms: *“If I were to take the survey and I’m experiencing worsening symptoms, I would definitely use the icons. This icon probably means that I can still function and things are tolerable in daily life.”* P7 similarly thought free-text annotations would help convey tolerability of withdrawal symptoms during the taper: *“If I had this app while I was tapering off my meds, I probably would have put a free-text note in there like, ‘I’m so dizzy that I’m unable to work.’”*

Participants sought to highlight various dimensions of their daily lives influenced by medication changes, such as work, social activities, or self-care. P4 desired to figure out how medication changes might impact his psychiatric symptoms in different social settings, which can be included in free-text annotations: *“I wonder how I’m going to feel, doing things outside my home, at work, social activities. So being able to address some of that would be helpful for the provider and myself.”* P8 wanted to specify what aspects of her daily life have been impacted by medication changes through free-text annotations: *“I feel like these [questions in the clinical survey] are just really broad. Some days, it was not an issue at all for me to take care of myself, but I couldn’t take care of social aspects of my life, and then other days, I couldn’t even get in the shower. So there needs to be an area to write specifically what that feeling was like.”*

5.1.2 Alleviating Cognitive and Emotional Burden of Logging. Participants recognized the potential of annotations in helping them engage in more patient-centered communication by alleviating the burden of logging. They envisioned that annotations would reduce the cognitive burden by highlighting primary concerns and helping recall symptoms, as well as the emotional burden associated with logging challenging mental health experiences.

Empowering the expression of primary concerns. Annotations were seen as a valuable means for participants to offload the cognitive burden of conveying their primary concerns with their providers. P7 wished to use free-text annotations to ensure that her concerns about withdrawal symptoms receive attention from her provider: *“If I had this app while I was tapering off my meds, I would have put a free-text note saying, ‘I’m so dizzy that I’m unable to work’ to make it clear to my provider that this is something that I’m really struggling with, and I need some changes to continue on my daily tasks.”* Visual annotations were also considered effective in conveying primary concerns. P7 believed these visual elements could highlight areas of significant impact: *“Amongst all of this data, it [an emoji] really highlights what I focused on most. I think the fact that I put an emoji there would show my psychiatrist that this issue is affecting me more heavily, whereas my average withdrawal symptom severity doesn’t have anything like that.”*

Enhancing recollection of symptom experiences. Participants found annotations valuable for reducing the cognitive burden of recalling their symptom experiences from the lengthy intervals between doctor’s visits. Given that most participants saw their providers only once or twice a year, they often felt that their progress was not adequately monitored during the tapering process. P7 expressed a desire for more frequent and in-depth discussions about medications: *“I wish that I had*

more frequent visits with him [provider], just to go over the medication a lot more in-depth. But we only talk three to five minutes at a time, and it's basically just so he can refill my prescription. I didn't really feel like my progress was being monitored." Annotations were thus seen as a potential way to bridge this gap by providing helpful references. P8 mentioned that free-text annotations would help her remember specific situations that influenced her mood: *"I'm not going to remember a few days later what I was specifically irritated or upset about. So this is going to be helpful."* P5 perceived that free-text annotations would help him not leave out important things about his symptom experiences during the doctor's visits: *"I always kind of worry, 'What if I left something out? What if I forget to say something in a session?' Having a report like this leaves what-if questions on the table because I've already done the reporting. It gives me more confidence in my general review of my symptoms."* P14 emphasized that visual annotations could trigger recollection of mental health symptoms during discussions with providers: *"If I add that [emoji], my provider can be like, 'Do you remember feeling like this?' It might be like a trigger to recall how I was feeling at that moment."*

Alleviating the emotional burden of logging. Visual annotations were perceived as a potentially powerful tool for alleviating the emotional burden associated with logging, allowing them to have more control over how they monitor their symptoms. Participants envisioned that when experiencing poor mental health, visual annotations would provide a more accessible and expressive means of conveying emotions. Some participants noted they might experience a lack of mental energy to write down how they were feeling, but selecting visual annotations that match their feelings would be much easier. P14 explained: *"If I'm in a bad state, I might not have the energy to fully write down how I'm feeling. I feel like an emoji or a GIF is gonna be a lot easier to relay at that time."* P16 similarly emphasized the ease of expressing emotions with visuals: *"Sometimes you don't feel like you can express it in words. So I think those [GIFs and icons] would be useful. I could look at some of these GIFs and choose Alice in Wonderland to express that feeling of getting overwhelmed by emotion or bewildered."* P19 also highlighted the value of visual annotations during times when verbal communication is difficult: *"Sometimes I don't feel like talking or even going to my appointment. If it's one of those days, I think that [visual annotations] would be a good way."*

Participants noted that visual annotations could help them document challenging mental health symptoms, which could be used later to convey those experiences to their providers. P18 explained that certain mental health experiences were overwhelming to verbalize or write about: *"Some of the stuff that is scarier to experience is also hard to verbalize. Writing it out makes it feel a lot more real and frightening."* The use of visual annotations alleviated some of this burden, allowing individuals to make a note for themselves without explicitly stating the details. P18 added: *"I do think that the icons are attractive. It is sort of a way that I can make a note for myself without having to deal with explicitly saying what it was, and then I can detail it later. If I'm verbalizing it at some time distance from the event, I think I would have an easier time describing it."*

5.2 Unmet Goals by Annotations

While participants perceived that annotations would help achieve the aforementioned communication goals in clinical settings, they felt annotations would fall short in (1) establishing a professional environment and (2) considering the sensitivity and complexity of mental health contexts.

5.2.1 Establishing a Professional Environment in Clinical Settings. Participants felt visual annotations might interrupt the professional relationship with providers by being too casual or ambiguous.

Too casual and light. Some participants felt that visual annotations often used in everyday interpersonal communication would not mesh with their professional communication environment in clinical settings because they might be too casual and light. P2 explained his perceptions about

emojis and GIFs: *“I feel like those silly GIFs are for adding comedy and levity. But depression and anxiety are not something that should be taken lightly. I feel like it [annotations] adds a sense of ‘It’s not that big of a deal’ and makes it cute and funny, so it just feels inappropriate.”* Several participants viewed annotations with emojis and GIFs as *“not very sincere”* (P16), *“superficial and flippant”* (P17), and for *“lightening up”* (P18), which led them to think those would not be appropriate for mental health contexts. P8 felt that such visual annotations might be particularly inappropriate if she was in poor mental health: *“If I was in an okay mood, I would use them [emojis, GIFs, icons]. Sometimes, I’d like to joke about things, so I can see how it can be a funny, humorous way to cope. But if I’m in such a bad mood, an emoji or a little picture is not gonna kind of sum that up.”*

Participants thus thought that, unlike in casual contexts, visual annotations might not be appropriate for communication with their providers. P18 said: *“I send those [emojis and GIFs] to try to lighten up a text to a friend. But this [visits for antidepressant management] doesn’t really feel like a lighten-it-up scenario.”* P5 similarly highlighted the inappropriateness of visual annotations in clinical settings: *“Emojis feel like interpersonal communication. It doesn’t generally mesh with the way that I communicate with a mental health professional.”* P19 perceived that visual annotations were for casual interactions on social media: *“It’s something that we’d use more on Facebook, and maybe that’s how the doctor is going to see it, very casual.”* They thus felt uncomfortable about using visual annotations for communication with their providers: *“Those [visual annotations] are more casual, and I would feel a little weird or uncomfortable about using a GIF in a note to a provider. (P15)”*

Too ambiguous and subjective. Some participants thought visual annotations might be too ambiguous and subjective, making it difficult for their providers to understand their intent. P11 was concerned about using visual annotations because the personal meanings may not be conveyed to her provider: *“I don’t think that those [emojis, icons, GIFs] would be helpful for providers. I know what I meant when I used these, but for a provider, it might not mean anything. They might be like, why would they add the Alice in Wonderland thing?”* P12 was similarly worried about her providers’ potential misunderstanding of her visual annotations: *“I would probably be less likely to use the GIFs with the doctor. Let’s say I put a GIF. They might interpret it differently from how I do.”*

5.2.2 Considering the Sensitivity and Complexity of Mental Health Contexts. Participants pointed out that annotations would fall short of considering the sensitivity and complexity of mental health contexts by requiring significant mental and emotional burdens and not providing the flexibility needed to convey complicated mental health experiences.

Mental and emotional burdens. Participants thought annotations might involve significant mental and emotional burdens for people when going through difficult mental health experiences. P1 stated she would not have the mental energy to add a free-text annotation to clinical surveys when in poor mental health: *“In the past when I was experiencing deep depression, I don’t know if I could properly articulate it. I can point to something real quick to convey how I’m feeling, but I don’t know whether I could add a note when things aren’t going well. For me, when depressed, ‘Go away. Don’t talk to me.’ I wanna put the covers over my head.”* P3 noted that visual annotations might require her too much mental energy when depressed: *“I feel like for someone who’s depressed, it [a GIF] is like brain overload. Like, ‘God, damn it! I gotta go through all of that?’ I’m exhausted just trying to choose which one.”* Participants also mentioned that they would feel more emotional burden when using personal language to log their mental health experiences through annotations. P18 said: *“I try to make things more objective and less emotional when I’m talking to my doctor, creating some distance between myself and my symptoms. When I started the medication, it was much easier to say, ‘I was suicidal’ than to talk about it in a more personal and emotional way. I am able to talk about more things if I’m able to put them in that [medical] language.”*

Inflexibility to convey complicated mental health experiences. Participants perceived that visual annotations would not be flexible enough to support them in conveying complicated mental health status. P5 thought visual annotations would not support as flexible communication as written words would do: *“I feel like emojis and GIFs are set in their nature. They’re not necessarily as flexible as the written word would be to me. They don’t cover the breadth of symptoms or emotions.”* P7 similarly felt: *“The emojis and GIFs can be used to convey emotion, but I think, with something so complicated, I would do better explaining how I really feel with written words. If I were to just pick one of the icons to represent how I’m feeling, I don’t think that would be sufficient.”*

6 Discussion

Our findings suggest that patients see annotations as a valuable tool for communicating their illness experiences, offering resources to convey a more authentic representation and influencing the direction of conversations during clinical visits. Our findings further show that patients consider various factors when selecting an annotation form, particularly how their providers will receive and interpret it. We recommend providing customization support to cater to patients’ communication needs, particularly considering provider perspectives. Lastly, it is crucial to consider the clinical practicality when designing annotation tools. We highlight opportunities to enhance the utility of digital symptom measure annotations in clinical settings.

6.1 Annotations For Facilitating Patient-Centered Communication

Through this study, we learned that patients often entered clinical visits with a clear idea about what they wanted to convey to their providers yet lacked the means to guide conversations around their complex symptom experiences when the interactions are centered around standardized patient self-reports. Aligned with prior work [6, 7, 28, 29, 32, 51, 82], this finding highlights the gap between patients’ desires to communicate their lived experiences and the limitations of current patient-provider communication practices. As the standardized self-report measures focus on the presence or absence of common symptoms, they often miss out on how the patient perceives and experiences those symptoms. Our findings suggest that annotations to clinical measures could provide patients with additional resources to describe their experiences in greater detail, potentially bridging communication gaps with providers. Our participants valued the space that annotations provide for documenting and elaborating on their experiences, envisioning that annotations could convey a more authentic representation of their illness experiences, such as symptom fluctuations, beyond being distilled down to *“just a number”* on a clinical scale. Our findings thus suggest that, unlike digital passive data collection methods, which are often held up as an alternative to self-report scales, annotations have the potential to offer patients a more authentic portrayal of patients’ lived experiences between visits while remaining grounded in standardized clinical measures.

Our findings further indicate the potential value of annotations in enhancing patient visibility in clinical settings. Our participants valued annotations for allowing them to convey *“why”* behind their clinical survey responses, such as their personal circumstances impacting mental health, and *how* the medication changes impacted their daily lives. Prior work suggested that solely relying on standardized measures could hinder patients’ sense of agency in managing their illness in everyday life, pointing to the need to provide mechanisms for patients to articulate and communicate personally meaningful aspects of their illness experiences [9, 10, 32, 82]. Our findings further highlight that the act of annotating clinical scales could provide patients with a sense of validation within the existing medical infrastructure, as their lived experiences are contextualized within a form that providers already engage with. This technique can empower patients through concrete means to have their voices heard and influence the direction of their communication with providers.

6.2 Considerations of the Form of Patient-Generated Data in Clinical Settings

Our study demonstrates that patients had particular forms of annotations that they felt comfortable presenting to their providers. We observed that patients considered various factors when selecting an annotation form, such as their mental health status, communication styles of themselves and their providers (e.g., level of formality), and the dynamics of patient-provider relationships (e.g., level of trust and bonding). For example, while some participants considered GIFs or emojis as an effective way of communicating their health experiences, others felt that such an approach lacked the necessary clarity for clinical settings. Further, while some participants felt GIFs and emojis could provide effective communication tools in clinical settings, others felt those forms were inappropriate for clinical contexts. These findings align with prior work, which highlighted patients' conflicting views of emojis as both lacking the seriousness needed for their challenging illness journey and as a valuable tool for reducing cognitive barriers in representing their personal health data by using symbols frequently used in everyday digital communication [97].

Factors such as patients' and providers' personalities, cultural or generational differences in experiences with different mediums, and perceived norms in patient-provider relationships may have influenced the contradictory viewpoints on annotation methods. In our study, participants who described themselves as more familiar with social media or messaging apps or were younger tended to prefer visual annotations, such as GIFs or emojis, over other forms like free-text notes, which aligns with previous research findings on the general use of GIFs outside health contexts [55, 67]. In contrast, patients who preferred more formal communication styles, were relatively less familiar with social media or messaging apps, or were older tended to be more reluctant to use GIFs or emojis and preferred data forms more common in clinical discussions, such as free-text notes and body parts. Preferences for different data forms are likely complex. For instance, some people might prefer not to use GIFs or emojis even if they are casual with their doctors because they do not regularly use these mediums in other communication in their daily lives.

We found it noteworthy that patients not only consider how helpful the data form will be for conveying their experiences but also how providers will receive and interpret it. While prior work on self-tracking has primarily focused on flexibility in data types to track more personally meaningful data [3, 9, 50, 51], our findings highlight that social settings warrant deeper consideration of the form and how other parties might interpret it. Studies on personal informatics in social media demonstrated that people often consider the form in which data are shared and how it would be perceived in social settings [38, 102]. Our findings further show that such consideration could be particularly critical in clinical contexts where care decisions could be made based on the tracked data. When designing a tool for supporting patient-centered communication through patient-generated data, it is crucial to recognize that the interpretation of data forms by other parties, particularly healthcare providers, can influence the effectiveness and reception of different data forms.

Future annotation tools could help people formally indicate their communication preferences and needs, particularly in clinical settings. Beyond presenting different styles of annotation forms, tools could allow people to explicitly specify communication preferences as part of aligning their annotations with their provider perspectives. Systems could then support customizing annotation strategies through the parameters indicating patients' communication needs. In particular, prior work highlighted that social contexts should be a primary consideration when designing for GIF and emoji use, suggesting opportunities to integrate GIF or emoji recommendations into keyboards personalized to each communication partner based on the chat log analysis [55, 58]. Using a similar approach, future tracking tools could provide recommendations for contents and data forms for their annotations based on patient input about their patient-provider dynamics. In addition, our study shows that patients often worry about the flexibility of GIFs and emojis to adequately describe

their complicated mental health status due to the limited options available online. One possible approach to address this concern is leveraging generative AI features, such as DALL-E [77], that can generate images from natural language descriptions to allow patients to quickly create visual artifacts and iterate on the outcome until it meets their articulation needs. Such a feature could further cater to aspects of patients' relationships with their providers (e.g., the level of formality).

However, given the sensitivity of patient-tracked data, privacy risk should be considered when implementing large language models and diffusion-infused features for annotation tools. For example, these tools likely need to be designed in line with medical data regulations, such as the Health Insurance Portability and Accountability Act (HIPAA) in the United States and the General Data Protection Regulation (GDPR) in the European Union. Incorporating ways to exclude or anonymize personally identifiable information from patient inputs could also help minimize privacy risks in using such features. More work is needed to explore the potential impact of tracking technology that supports the customization of annotations to clinical measures.

6.3 Balancing Patient-Centric Communication with Clinical Practicality

Our study highlights the potential of annotations to clinical symptom measures to enhance patient-centered communication. Grounded in clinical symptom measures, patients felt that annotations could help them convey their lived experiences to their providers while maintaining clinical relevance. Nevertheless, it is crucial to consider the practical aspects of implementing a tool for supporting annotations in real-world clinical contexts. Prior work suggested that patient-generated, free-form data may not always be well-received by providers, particularly because patients tend to bring in more data than what providers can realistically review within the time constraints of clinical visits [49, 59, 91, 93, 94, 96, 105, 107]. In our study, some participants also mentioned that their typical visits were as short as under five minutes, which suggests it might be infeasible or inadequate to include additional artifacts or processes. Further, participants were worried that their annotations could come across as overwhelming to their providers, making them skeptical about whether their providers would actually review their annotations in practice. If annotations are not reviewed or discussed, patients could potentially feel that their sense of agency is undermined even more than if they only completed a standard self-report scale. Therefore, in the design of digital symptom measure tools, it is critical to consider ways to balance patient-centric communication with the practical demands of clinical settings.

To enhance the utility of annotations to self-report measures in clinical settings, a digital symptom measure annotation tool could aid patients by offering guidance about the types of information that providers might find useful to include in an annotation and explaining why this information could be helpful. After patients create some annotations, tools could assist patients in reflecting on what annotations could be most informational to a provider, such as data that are indicative of worsening symptoms. In our study, participants often felt that the annotations could be helpful not only for provider review but also for self-reflection, which suggests the potential value of annotations for both purposes. In light of this finding, annotation tools could suggest keeping annotations that are less clinically relevant as personal journals, such as daily mood fluctuations, even if the annotations do not show up for provider review. Further, tools could automatically generate a bullet-point summary from these annotations, providing patients with tailored recommendations on critical points to discuss in their next clinical visit. Such approaches could help ensure that the annotations presented and discussed at clinical visits are directly pertinent to the patient's ongoing care, making patient-driven self-tracking data more practical for care processes.

7 Conclusion

Through interviews with 20 patients who were either interested in stopping their antidepressants or in the process of discontinuing them, we found that annotations to clinical measures were seen as a useful tool to enrich clinical symptom measures with individual illness experiences and alleviate the cognitive and emotional burden of logging. However, patients also thought annotations might interrupt the professional relationship with their providers and overlook the sensitivity and complexity of mental health contexts. Based on the findings, we suggest opportunities for annotations to promote patient-centered communication. We further propose incorporating customization support for patients' communication needs around the form of patient-generated data. Lastly, we highlight the need to develop ways to enhance the clinical practicality of annotations.

Acknowledgments

We are grateful to our participants for providing valuable insights and to Bryan Shapiro for assisting with recruitment. We thank Alison Liu and Charlotte Moremen for their support in designing a prototype used in the study and conducting interviews. This project was partially supported by the National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health, through grant UL1 TR001414 as well as the National Science Foundation under Award IIS-2237389. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the National Science Foundation.

References

- [1] Substance Abuse and Mental Health Services Administration. 2019. Key substance use and mental health indicators in the United States: Results from the 2018 National Survey on Drug Use and Health. *HHS Publication No. PEP19-5068, NSDUH Series H-54* 170 (2019), 51–58. <https://www.samhsa.gov/data/>
- [2] Alexander T. Adams, Elizabeth L. Murnane, Phil Adams, Michael Elfenbein, Pamara F. Chang, Shruti Sannon, Geri Gay, and Tanzeem Choudhury. 2018. Keppi: A Tangible User Interface for Self-Reporting Pain. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (CHI '18). Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3173574.3174076>
- [3] Phil Adams, Elizabeth L. Murnane, Michael Elfenbein, Elaine Wethington, and Geri Gay. 2017. Supporting the Self-Management of Chronic Pain Conditions with Tailored Momentary Self-Assessments. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) (CHI '17). Association for Computing Machinery, New York, NY, USA, 1065–1077. <https://doi.org/10.1145/3025453.3025832>
- [4] Maayan Agmon, Anna Zisberg, Efrat Gil, Debbie Rand, Nurit Gur-Yaish, and Mary Azriel. 2017. Adult Utilization of Psychiatric Drugs and Differences by Sex, Age, and Race. *JAMA Internal Medicine* 177, 2 (2017), 272–274. <https://doi.org/10.1001/jamainternmed.2016.7266>
- [5] Tariq Andersen, Jørgen Bansler, Finn Kensing, Jonas Moll, and Karen Dam Nielsen. 2014. Alignment of concerns: A design rationale for patient participation in eHealth. *Proceedings of the Annual Hawaii International Conference on System Sciences* (2014), 2587–2596. <https://doi.org/10.1109/HICSS.2014.327>
- [6] Tariq Andersen, Pernille Bjørn, Finn Kensing, and Jonas Moll. 2011. Designing for collaborative interpretation in telemonitoring: Re-introducing patients as diagnostic agents. *International Journal of Medical Informatics* 80, 8 (2011), e112–e126. <https://doi.org/10.1016/j.ijmedinf.2010.09.010>
- [7] Tariq Osman Andersen, Jørgen Peter Bansler, Finn Kensing, Jonas Moll, Troels Mønsted, Karen Dam Nielsen, Olav Wendelboe Nielsen, Helen Høgh Petersen, and Jesper Hastrup Svendsen. 2019. Aligning concerns in telecare: three concepts to guide the design of patient-centred E-health. *Computer Supported Cooperative Work (CSCW)* 28 (2019), 1039–1072. <https://doi.org/10.1007/s10606-018-9309-1>
- [8] American Psychiatric Association. 2010. Treatment of Patients With Major Depressive Disorder Second Edition APA. *Psychiatric Services* April (2010), 1–78.
- [9] Amid Ayobi, Paul Marshall, and Anna L. Cox. 2020. Trackly: A Customisable and Pictorial Self-Tracking App to Support Agency in Multiple Sclerosis Self-Care. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–15. <https://doi.org/10.1145/3313831.3376809>

- [10] Amid Ayobi, Paul Marshall, Anna L. Cox, and Yunan Chen. 2017. Quantifying the Body and Caring for the Mind: Self-Tracking in Multiple Sclerosis. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) (CHI '17). Association for Computing Machinery, New York, NY, USA, 6889–6901. <https://doi.org/10.1145/3025453.3025869>
- [11] Amid Ayobi, Tobias Sonne, Paul Marshall, and Anna L. Cox. 2018. Flexible and Mindful Self-Tracking: Design Implications from Paper Bullet Journals. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (CHI '18). Association for Computing Machinery, New York, NY, USA, 1–14. <https://doi.org/10.1145/3173574.3173602>
- [12] Saeideh Bakhshi, David A. Shamma, Lyndon Kennedy, Yale Song, Paloma de Juan, and Joseph 'Jofish' Kaye. 2016. Fast, Cheap, and Good: Why Animated GIFs Engage Us. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (CHI '16). Association for Computing Machinery, New York, NY, USA, 575–586. <https://doi.org/10.1145/2858036.2858532>
- [13] Jakob E. Bardram, Mads Frost, Károly Szántó, Maria Faurholt-Jepsen, Maj Vinberg, and Lars Vedel Kessing. 2013. Designing mobile health technology for bipolar disorder: a field trial of the monarca system. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) (CHI '13). Association for Computing Machinery, New York, NY, USA, 2627–2636. <https://doi.org/10.1145/2470654.2481364>
- [14] A. T. Beck. 1961. An Inventory for Measuring Depression. *Archives of General Psychiatry* 4, 6 (jun 1961), 561. <https://doi.org/10.1001/archpsyc.1961.01710120031004>
- [15] Andrew B.L. Berry, Catherine Lim, Andrea L. Hartzler, Tad Hirsch, Evette Ludman, Edward H. Wagner, and James D. Ralston. 2017. Creating Conditions for Patients' Values to Emerge in Clinical Conversations: Perspectives of Health Care Team Members. In *Proceedings of the 2017 Conference on Designing Interactive Systems* (Edinburgh, United Kingdom) (DIS '17). Association for Computing Machinery, New York, NY, USA, 1165–1174. <https://doi.org/10.1145/3064663.3064669>
- [16] Andrew BL Berry, Catherine Lim, Andrea L Hartzler, Tad Hirsch, Evette Ludman, Edward H Wagner, and James D Ralston. 2017. Eliciting values of patients with multiple chronic conditions: Evaluation of a patient-centered framework. In *AMIA Annual Symposium Proceedings*, Vol. 2017. American Medical Informatics Association, 430.
- [17] Andrew B.L. Berry, Catherine Y. Lim, Andrea L. Hartzler, Tad Hirsch, Evette Ludman, Edward H. Wagner, and James D. Ralston. 2017. "It's good to know you're not a stranger every time": Communication about values between patients with multiple chronic conditions and healthcare providers. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017). <https://doi.org/10.1145/3134658>
- [18] Andrew B.L. Berry, Catherine Y. Lim, Calvin A. Liang, Andrea L. Hartzler, Tad Hirsch, Dawn M. Ferguson, Zoë A. Bermet, and James D. Ralston. 2021. Supporting Collaborative Reflection on Personal Values and Health. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021). <https://doi.org/10.1145/3476040>
- [19] Andrew B. L. Berry, Catherine Y. Lim, Tad Hirsch, Andrea L. Hartzler, Linda M. Kiel, Zoë A. Bermet, and James D. Ralston. 2019. Supporting Communication About Values Between People with Multiple Chronic Conditions and their Providers. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland UK) (CHI '19). Association for Computing Machinery, New York, NY, USA, 1–14. <https://doi.org/10.1145/3290605.3300700>
- [20] Nick Black. 2013. Patient reported outcome measures could help transform healthcare. *BMJ (Online)* 346, 7896 (2013), 1–5. <https://doi.org/10.1136/bmj.f167>
- [21] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101.
- [22] Patricia Flatley Brennan and Gail Casper. 2015. Observing health in everyday living: ODLs and the care-between-the-care. *Personal and Ubiquitous Computing* 19, 1 (2015), 3–8. <https://doi.org/10.1007/s00779-014-0805-0>
- [23] Debra J. Brody and Qiuping Gu. 2020. Antidepressant Use Among Adults: United States, 2015–2018. *NCHS data brief* 377 (2020), 1–8.
- [24] Michael Brundage, Jane Blazeby, Dennis Revicki, Brenda Bass, Henrica De Vet, Helen Duffy, Fabio Efficace, Madeleine King, Cindy L.K. Lam, David Moher, Jane Scott, Jeff Sloan, Claire Snyder, Susan Yount, and Melanie Calvert. 2013. Patient-reported outcomes in randomized clinical trials: Development of ISOQOL reporting standards. *Quality of Life Research* 22, 6 (2013), 1161–1175. <https://doi.org/10.1007/s11136-012-0252-1>
- [25] Scott A. Bull, Enid M. Hunkeler, Janelle Y. Lee, Clayton R. Rowland, Todd E. Williamson, Joseph R. Schwab, Stephen W. Hurt, Lydia González, and Denyse Demers. 2002. Discontinuing or switching selective serotonin-reuptake inhibitors. *Annals of Pharmacotherapy* 36, 4 (2002), 578–584. <https://doi.org/10.1345/aph.1A254>
- [26] Ayşe G. Büyüktür and Mark S. Ackerman. 2017. Information Work in Bone Marrow Transplant: Reducing Misalignment of Perspectives. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, Oregon, USA) (CSCW '17). Association for Computing Machinery, New York, NY, USA, 1740–1752. <https://doi.org/10.1145/2998181.2998361>

- [27] Arlene E. Chung and Ethan M. Basch. 2015. Incorporating the patient's voice into electronic health records through patient-reported outcomes as the "review of systems". *Journal of the American Medical Informatics Association* 22, 4 (2015), 914–916. <https://doi.org/10.1093/jamia/ocu007>
- [28] Chia-Fang Chung, Kristin Dew, Allison Cole, Jasmine Zia, James Fogarty, Julie A. Kientz, and Sean A. Munson. 2016. Boundary Negotiating Artifacts in Personal Informatics: Patient-Provider Collaboration with Patient-Generated Data. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (San Francisco, California, USA) (CSCW '16). Association for Computing Machinery, New York, NY, USA, 770–786. <https://doi.org/10.1145/2818048.2819926>
- [29] Chia-Fang Chung, Qiaosi Wang, Jessica Schroeder, Allison Cole, Jasmine Zia, James Fogarty, and Sean A. Munson. 2019. Identifying and Planning for Individualized Change: Patient-Provider Collaboration Using Lightweight Food Diaries in Healthy Eating and Irritable Bowel Syndrome. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 1, Article 7 (mar 2019), 27 pages. <https://doi.org/10.1145/3314394>
- [30] Charles S. Cleeland and Katelyn M. Ryan. 1994. Pain assessment: global use of the Brief Pain Inventory. *Annals of the Academy of Medicine, Singapore* 23 2 (1994), 129–38.
- [31] John J. Collins, Maura E. Byrnes, Ira J. Dunkel, Jeanne Lapin, Traci Nadel, Howard T. Thaler, Tanya Polyak, Bruce Rapkin, and Russell K. Portenoy. 2000. The measurement of symptoms in children with cancer. *Journal of Pain and Symptom Management* 19, 5 (2000), 363–377. [https://doi.org/10.1016/S0885-3924\(00\)00127-5](https://doi.org/10.1016/S0885-3924(00)00127-5)
- [32] Mayara Costa Figueiredo, H. Irene Su, and Yunan Chen. 2021. Using Data to Approach the Unknown: Patients' and Healthcare Providers? Data Practices in Fertility Challenges. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW3 (Jan 2021), 1–35. <https://doi.org/10.1145/3432926>
- [33] Henriette Cramer, Paloma De Juan, and Joel Tetreault. 2016. Sender-intended functions of emojis in US messaging. *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI 2016* (2016), 504–509. <https://doi.org/10.1145/2935334.2935370>
- [34] James Davies and John Read. 2019. A systematic review into the incidence, severity and duration of antidepressant withdrawal effects: Are guidelines evidence-based? *Addictive Behaviors* 97, September 2018 (2019), 111–121. <https://doi.org/10.1016/j.addbeh.2018.08.027>
- [35] Prasanna R. Deshpande, BLakshmi Sudeepthi, Surulivel Rajan, and CP Abdul Nazir. 2011. Patient-reported outcomes: A new era in clinical research. *Perspectives in Clinical Research* 2, 4 (2011), 137. <https://doi.org/10.4103/2229-3485.86879>
- [36] K J Devers. 1999. How will we know "good" qualitative research when we see it? Beginning the dialogue in health services research. *Health services research* 34, 5 Pt 2 (1999), 1153–88. <http://www.ncbi.nlm.nih.gov/pubmed/10591278%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1089058>
- [37] Astrid Eliassen, Mikkel Kramme Abildtoft, Niels Steen Krogh, Catherine Rechnitzer, Jesper Sune Brok, René Mathiasen, Kjeld Schmiegelow, and Kim Peder Dalhoff. 2020. Smartphone app to self-monitor nausea during pediatric chemotherapy treatment: User-centered design process. *JMIR mHealth and uHealth* 8, 7 (2020), 1–11. <https://doi.org/10.2196/18564>
- [38] Daniel A. Epstein, Bradley H. Jacobson, Elizabeth Bales, David W. McDonald, and Sean A. Munson. 2015. From "nobody cares" to "way to go!": A Design Framework for Social Sharing in Personal Informatics. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (Vancouver, BC, Canada) (CSCW '15). Association for Computing Machinery, New York, NY, USA, 1622–1636. <https://doi.org/10.1145/2675133.2675135>
- [39] Nigel G Fielding. 2011. Self-Report Study: The SAGE Dictionary of Social Research Methods. *The SAGE Dictionary of Social Research Methods* (2011), 276–277.
- [40] Mayara Costa Figueiredo and Yunan Chen. 2020. Patient-generated health data: Dimensions, challenges, and open questions. *Foundations and Trends in Human-Computer Interaction* 13, 3 (2020), 165–297. <https://doi.org/10.1561/1100000080>
- [41] National Institute for Health Care and Excellence. 2018. Depression in adults: treatment and management Full guideline. May (2018).
- [42] Adele Framer. 2021. What I have learnt from helping thousands of people taper off antidepressants and other psychotropic medications. *Therapeutic Advances in Psychopharmacology* 11 (2021), 204512532199127. <https://doi.org/10.1177/2045125321991274>
- [43] Mads Frost, Afsaneh Doryab, Maria Faurholt-Jepsen, Lars Vedel Kessing, and Jakob E. Bardram. 2013. Supporting disease insight through data analysis: refinements of the monarca self-assessment system. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Zurich, Switzerland) (UbiComp '13). Association for Computing Machinery, New York, NY, USA, 133–142. <https://doi.org/10.1145/2493432.2493507>
- [44] Erik Grönvall and Nervo Verdezoto. 2013. Beyond self-monitoring: Understanding non-functional aspects of home-based healthcare technology. *UbiComp 2013 - Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (2013), 587–596. <https://doi.org/10.1145/2493432.2493495>

- [45] Peter C. Groot and Jim van Os. 2020. How user knowledge of psychotropic drug withdrawal resulted in the development of person-specific tapering medication. *Therapeutic Advances in Psychopharmacology* 10 (2020), 204512532093245. <https://doi.org/10.1177/2045125320932452>
- [46] Peter Haddad. 1998. The SSRI discontinuation syndrome. *Journal of psychopharmacology* 12, 3 (1998), 305–313.
- [47] M. Hamilton. 1960. A RATING SCALE FOR DEPRESSION. *Journal of Neurology, Neurosurgery & Psychiatry* 23, 1 (feb 1960), 56–62. <https://doi.org/10.1136/jnnp.23.1.56>
- [48] Andrea Hartzler and Wanda Pratt. 2011. Managing the Personal Side of Health: How Patient Expertise Differs from the Expertise of Clinicians. *Journal of Medical Internet Research* 13, 3 (Aug. 2011), e62. <https://doi.org/10.2196/jmir.1728>
- [49] Michael Jeffrey Daniel Hoefer, Lucy Van Kleunen, Cassandra Goodby, Lanea Blyss Blackburn, Priyanka Panati, and Stephen Volda. 2021. The Multiplicative Patient and the Clinical Workflow: Clinician Perspectives on Social Interfaces for Self-Tracking and Managing Bipolar Disorder. In *Proceedings of the 2021 ACM Designing Interactive Systems Conference (Virtual Event, USA) (DIS '21)*. Association for Computing Machinery, New York, NY, USA, 907–925. <https://doi.org/10.1145/3461778.3461995>
- [50] Matthew K. Hong, Udaya Lakshmi, Kimberly Do, Sampath Prahalad, Thomas Olson, Rosa I. Arriaga, and Lauren Wilcox. 2020. Using Diaries to Probe the Illness Experiences of Adolescent Patients and Parental Caregivers. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (Honolulu, HI, USA) (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–16. <https://doi.org/10.1145/3313831.3376426>
- [51] Matthew K. Hong, Udaya Lakshmi, Thomas A. Olson, and Lauren Wilcox. 2018. Visual ODLs: Co-Designing Patient-Generated Observations of Daily Living to Support Data-Driven Conversations in Pediatric Care. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (Montreal QC, Canada) (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3173574.3174050>
- [52] Mark Abie Horowitz and David Taylor. 2019. Tapering of SSRI treatment to mitigate withdrawal symptoms. *The Lancet Psychiatry* 6, 6 (2019), 538–546. [https://doi.org/10.1016/s2215-0366\(19\)30032-x](https://doi.org/10.1016/s2215-0366(19)30032-x)
- [53] Maia L. Jacobs, James Clawson, and Elizabeth D. Mynatt. 2015. Comparing Health Information Sharing Preferences of Cancer Patients, Doctors, and Navigators. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (Vancouver, BC, Canada) (CSCW '15)*. Association for Computing Machinery, New York, NY, USA, 808–818. <https://doi.org/10.1145/2675133.2675252>
- [54] Amy Jang, Diana L. MacLean, and Jeffrey Heer. 2014. BodyDiagrams: improving communication of pain symptoms through drawing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Toronto, Ontario, Canada) (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 1153–1162. <https://doi.org/10.1145/2556288.2557223>
- [55] Jialun "Aaron" Jiang, Casey Fiesler, and Jed R. Brubaker. 2018. 'The Perfect One': Understanding Communication Practices and Challenges with Animated GIFs. *Proceedings of the ACM on human-computer interaction* 2, CSCW, Article 80 (Nov 2018), 20 pages. <https://doi.org/10.1145/3274349>
- [56] Eunkyung Jo, Myeonghan Ryu, Georgia Kenderova, Samuel So, Bryan Shapiro, Alexandra Papoutsaki, and Daniel A. Epstein. 2022. Designing Flexible Longitudinal Regimens: Supporting Clinician Planning for Discontinuation of Psychiatric Drugs. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (New Orleans, LA, USA) (CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 352, 17 pages. <https://doi.org/10.1145/3491102.3502206>
- [57] Ravi Karkar, Jessica Schroeder, Daniel A. Epstein, Laura R. Pina, Jeffrey Scofield, James Fogarty, Julie A. Kientz, Sean A. Munson, Roger Vilardaga, and Jasmine Zia. 2017. TummyTrials: A Feasibility Study of Using Self-Experimentation to Detect Individualized Food Triggers. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (Denver, Colorado, USA) (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 6850–6863. <https://doi.org/10.1145/3025453.3025480>
- [58] Joongyung Kim, Taesik Gong, Bogoan Kim, Jaeyeon Park, Woojeong Kim, Evey Huang, Kyungsik Han, Juho Kim, Jeonggil Ko, and Sung-Ju Lee. 2020. No More One Liners. *ACM Transactions on Social Computing* 3, 2 (2020), 1–25. <https://doi.org/10.1145/3373146>
- [59] Yoojung Kim, Eunyoung Heo, Hyunjeong Lee, Sookyoung Ji, Jueun Choi, Jeong-Whun Kim, Joongseek Lee, and Sooyoung Yoo. 2017. Prescribing 10,000 Steps Like Aspirin: Designing a Novel Interface for Data-Driven Medical Consultations. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (Denver, Colorado, USA) (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 5787–5799. <https://doi.org/10.1145/3025453.3025570>
- [60] Yoojung Kim, Sookyoung Ji, Hyunjeong Lee, Jeong-Whun Kim, Sooyoung Yoo, and Joongseek Lee. 2016. "My Doctor is Keeping an Eye on Me!": Exploring the Clinical Applicability of a Mobile Food Logger. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (San Jose, California, USA) (CHI '16)*. Association for Computing Machinery, New York, NY, USA, 5620–5631. <https://doi.org/10.1145/2858036.2858145>

- [61] Kurt Kroenke, Robert L. Spitzer, and Janet B. W. Williams. 2001. The PHQ-9: Validity of of a Brief Depression Severity Measure. *Journal of General Internal Medicine* 16, 9 (sep 2001), 606–613. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x> arXiv:arXiv:1011.1669v3
- [62] Udaya Lakshmi, Matthew Hong, and Lauren Wilcox. 2018. Integrating Patient-Generated Observations of Daily Living into Pediatric Cancer Care: A Formative User Interface Design Study. In *2018 IEEE International Conference on Healthcare Informatics (ICHI)*. 265–275. <https://doi.org/10.1109/ICHI.2018.00037>
- [63] Charlotte P. Lee. 2007. Boundary negotiating artifacts: Unbinding the routine of boundary objects and embracing chaos in collaborative work. *Computer Supported Cooperative Work* 16, 3 (2007), 307–339. <https://doi.org/10.1007/s10606-007-9044-5>
- [64] Catherine Lim, Andrew B.L. Berry, Tad Hirsch, Andrea L. Hartzler, Edward H. Wagner, Evette Ludman, and James D. Ralston. 2016. "It just seems outside my health": How Patients with Chronic Conditions Perceive Communication Boundaries with Providers. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems* (Brisbane, QLD, Australia) (DIS '16). Association for Computing Machinery, New York, NY, USA, 1172–1184. <https://doi.org/10.1145/2901790.2901866>
- [65] Catherine Y. Lim, Andrew B.L. Berry, Andrea L. Hartzler, Tad Hirsch, David S. Carrell, Zoë A. Bermet, and James D. Ralston. 2019. Facilitating Self-reflection about Values and Self-care Among Individuals with Chronic Conditions. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) (CHI '19). Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/3290605.3300885>
- [66] Sarah E. Lord, Kimberlee J. Trudeau, Ryan A. Black, Lucy Lorin, Elizabeth Cooney, Albert Villapiano, and Stephen F. Butler. 2011. CHAT: Development and validation of a computer-delivered, self-report, substance use assessment for adolescents. *Substance Use and Misuse* 46, 6 (2011), 781–794. <https://doi.org/10.3109/10826084.2010.538119>
- [67] Xuan Lu, Wei Ai, Xuanzhe Liu, Qian Li, Ning Wang, Gang Huang, and Qiaozhu Mei. 2016. Learning from the ubiquitous language: an empirical analysis of emoji usage of smartphone users. , 11 pages. <https://doi.org/10.1145/2971648.2971724>
- [68] Yuhan Luo, Young Ho Kim, Bongshin Lee, Naeemul Hassan, and Eun Kyoung Choe. 2021. FoodScrap: Promoting Rich Data Capture and Reflective Food Journaling through Speech Input. *DIS 2021 - Proceedings of the 2021 ACM Designing Interactive Systems Conference: Nowhere and Everywhere* (2021), 606–618. <https://doi.org/10.1145/3461778.3462074>
- [69] Courtney R. Lyles, Andrea Altschuler, Neetu Chawla, Christine Kowalski, Deanna McQuillan, Elizabeth Bayliss, Michele Heisler, and Richard W. Grant. 2016. User-centered design of a tablet waiting room tool for complex patients to prioritize discussion topics for primary care visits. *JMIR mHealth and uHealth* 4, 3 (2016), 1–10. <https://doi.org/10.2196/mhealth.6187>
- [70] Mark Matthews, Elizabeth Murnane, and Jaime Snyder. 2017. Quantifying the Changeable Self: The Role of Self-Tracking in Coming to Terms With and Managing Bipolar Disorder. *Human-Computer Interaction* 32, 5-6 (2017), 413–446. <https://doi.org/10.1080/07370024.2017.1294983>
- [71] Mark Matthews, Stephen Volda, Saeed Abdullah, Gavin Doherty, Tanzeem Choudhury, Sangha Im, and Geri Gay. 2015. In situ design for mental illness: Considering the pathology of bipolar disorder in mhealth design. *MobileHCI 2015 - Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services* (2015), 86–97. <https://doi.org/10.1145/2785830.2785866>
- [72] Sonali R. Mishra, Predrag Klasnja, John MacDuffie Woodburn, Eric B. Hekler, Larsson Omberg, Michael Kellen, and Lara Mangravite. 2019. Supporting Coping with Parkinson's Disease Through Self Tracking. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–16. <https://doi.org/10.1145/3290605.3300337>
- [73] Sonali R. Mishra, Andrew D. Miller, Shefali Halder, Maher Khelifi, Jordan Eschler, Rashmi G. Elera, Ari H. Pollack, and Wanda Pratt. 2018. Supporting Collaborative Health Tracking in the Hospital: Patients' Perspectives. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (CHI '18). Association for Computing Machinery, New York, NY, USA, 1–14. <https://doi.org/10.1145/3173574.3174224>
- [74] Elizabeth L. Murnane, Dan Cosley, Pamara Chang, Shion Guha, Ellen Frank, Geri Gay, and Mark Matthews. 2016. Self-monitoring practices, attitudes, and needs of individuals with bipolar disorder: Implications for the design of technologies to manage mental health. *Journal of the American Medical Informatics Association* 23, 3 (2016), 477–484. <https://doi.org/10.1093/jamia/ocv165>
- [75] Royal College of Psychiatrists. 2019. Position statement on antidepressants and depression. https://www.rcpsych.ac.uk/docs/default-source/improving-care/better-mh-policy/position-statements/ps04_19---antidepressants-and-depression.pdf?sfvrsn=ddea9473_5
- [76] Nikki R. Ogle and Shawn R. Akkerman. 2013. Guidance for the discontinuation or switching of antidepressant therapies in adults. *Journal of Pharmacy Practice* 26, 4 (2013), 389–396. <https://doi.org/10.1177/0897190012467210>
- [77] OpenAI, Inc. 2023. DALL-E-3. Retrieved Dec 14, 2023 from <https://openai.com/dall-e-3>

- [78] Jasper E. Palmier-Claus, John Ainsworth, Matthew Machin, Cristine Barrowclough, Graham Dunn, Emma Barkus, Anne Rogers, Til Wykes, Shitij Kapur, Iain Buchan, Emma Salter, and Shôn W. Lewis. 2012. The feasibility and validity of ambulatory self-report of psychotic symptoms using a smartphone software application. *BMC Psychiatry* 12, 1 (2012), 1. <https://doi.org/10.1186/1471-244X-12-172>
- [79] Alexandra Papoutsaki, Samuel So, Georgia Kenderova, Bryan Shapiro, and Daniel A. Epstein. 2021. Understanding Delivery of Collectively Built Protocols in an Online Health Community for Discontinuation of Psychiatric Drugs. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 420 (oct 2021), 29 pages. <https://doi.org/10.1145/3479564>
- [80] Rupa A Patel, Predrag Klasnja, Andrea Hartzler, Kenton T Unruh, and Wanda Pratt. 2012. Probing the benefits of real-time tracking during cancer care. In *AMIA Annual Symposium Proceedings*, Vol. 2012. American Medical Informatics Association, 1340.
- [81] Ryne Paulose-Ram, Marc A. Safran, Bruce S. Jonas, Qiuping Gu, and Denise Orwig. 2007. Trends in psychotropic medication use among U.S. adults. *Pharmacoepidemiology and Drug Safety* 16, 5 (may 2007), 560–570. <https://doi.org/10.1002/pds.1367>
- [82] Adrienne Pichon, Kayla Schiffer, Emma Horan, Bria Massey, Suzanne Bakken, Lena Mamykina, and Noemie Elhadad. 2021. Divided We Stand: The Collaborative Work of Patients and Providers in an Enigmatic Chronic Disease. *Proceedings of the ACM on human-computer interaction* 4, CSCW3, Article 261 (2021), 24 pages. <https://doi.org/10.1145/3434170>
- [83] Blaine A. Price, Ryan Kelly, Vikram Mehta, Ciaran McCormick, Hanad Ahmed, and Oliver Pearce. 2018. Feel My Pain: Design and Evaluation of Painpad, a Tangible Device for Supporting Inpatient Self-Logging of Pain. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (CHI '18). Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3173574.3173743>
- [84] Shriti Raj, Mark W. Newman, Joyce M. Lee, and Mark S. Ackerman. 2017. Understanding individual and collaborative problem-solving with patient-generated data: Challenges and opportunities. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017), 1–18. <https://doi.org/10.1145/3134723>
- [85] John Read. 2020. How common and severe are six withdrawal effects from, and addiction to, antidepressants? The experiences of a large international sample of patients. *Addictive Behaviors* 102, July 2019 (2020), 106157. <https://doi.org/10.1016/j.addbeh.2019.106157>
- [86] John Read, Claire Cartwright, and Kerry Gibson. 2018. How many of 1829 antidepressant users report withdrawal effects or addiction? *International Journal of Mental Health Nursing* 27, 6 (2018), 1805–1815. <https://doi.org/10.1111/inm.12488>
- [87] Fujiko Robledo Yamamoto, Amy Volda, and Stephen Volda. 2021. From Therapy to Teletherapy: Relocating Mental Health Services Online. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021). <https://doi.org/10.1145/3479508>
- [88] Jerrold F. Rosenbaum, Maurizio Fava, Sharon L. Hoog, Richard C. Ascroft, and William B. Krebs. 1998. Selective serotonin reuptake inhibitor discontinuation syndrome: A randomized clinical trial. *Biological Psychiatry* 44, 2 (1998), 77–87. [https://doi.org/10.1016/S0006-3223\(98\)00126-7](https://doi.org/10.1016/S0006-3223(98)00126-7)
- [89] Edgar L. Ross, Robert N. Jamison, Lance Nicholls, Barbara M. Perry, and Kim D. Nolen. 2020. Clinical integration of a smartphone app for patients with chronic pain: Retrospective analysis of predictors of benefits and patient engagement between clinic visits. *Journal of Medical Internet Research* 22, 4 (2020), 1–13. <https://doi.org/10.2196/16939>
- [90] Johnny Saldaña. 2022. *The coding manual for qualitative researchers*. Sage.
- [91] Patrick C. Sanger, Andrea Hartzler, Ross J. Lordon, Cheryl AL Armstrong, William B. Lober, Heather L. Evans, and Wanda Pratt. 2016. A patient-centered system in a provider-centered world: challenges of incorporating post-discharge wound data into practice. *Journal of the American Medical Informatics Association* 23, 3 (may 2016), 514–525. <https://doi.org/10.1093/jamia/ocv183>
- [92] Maria J. Santana, Lotte Haverman, Kate Absolom, Elena Takeuchi, David Feeny, Martha Grootenhuis, and Galina Velikova. 2015. Training clinicians in how to use patient-reported outcome measures in routine clinical practice. *Quality of Life Research* 24, 7 (2015), 1707–1718. <https://doi.org/10.1007/s11136-014-0903-5>
- [93] Jessica Schroeder, Jane Hoffswell, Chia Fang Chung, James Fogarty, Sean Munson, and Jasmine Zia. 2017. Supporting patient-provider collaboration to identify individual triggers using food and symptom journals. *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, CSCW (2017), 1726–1739. <https://doi.org/10.1145/2998181.2998276>
- [94] Jessica Schroeder, Ravi Karkar, Natalia Murinova, James Fogarty, and Sean A. Munson. 2019. Examining opportunities for goal-directed self-tracking to support chronic condition management. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 4 (2019). <https://doi.org/10.1145/3369809>
- [95] Gro Hilde Severinsen, Line Silsand, Gunnar Ellingsen, and Rune Pedersen. 2019. From Free-Text to Structure in Electronic Patient Records. *Studies in Health Technology and Informatics* 265 (2019), 86–91. <https://doi.org/10.3233/SHIT190143>

- [96] Hardeep Singh, Farah Tahsin, Jason Xin Nie, Brian McKinstry, Kednapa Thavorn, Ross Upshur, Sarah Harvey, Walter P. Wodchis, and Carolyn Steele Gray. 2021. Exploring the perspectives of primary care providers on use of the electronic Patient Reported Outcomes tool to support goal-oriented care: a qualitative study. *BMC Medical Informatics and Decision Making* 21, 1 (2021), 1–15. <https://doi.org/10.1186/s12911-021-01734-0>
- [97] Lauren E. Snyder, Daniel F. Phan, Kristen C. Williams, Eduardo Piqueiras, Sarah E. Connor, Sheba George, Lorna Kwan, Jefersson Villatoro Chavez, Megha D. Tandel, Stanley K. Frencher, Mark S. Litwin, John L. Gore, and Andrea L. Hartzler. 2022. Comprehension, utility, and preferences of prostate cancer survivors for visual timelines of patient-reported outcomes co-designed for limited graph literacy: Meters and emojis over comics. *Journal of the American Medical Informatics Association* 29, 11 (2022), 1838–1846. <https://doi.org/10.1093/jamia/ocac148>
- [98] Simone Spark, Dyani Lewis, Alaina Vaisey, Eris Smyth, Anna Wood, Meredith Temple-Smith, Rebecca Lorch, Rebecca Guy, and Jane Hocking. 2015. Using computer-assisted survey instruments instead of paper and pencil increased completeness of self-administered sexual behavior questionnaires. *Journal of Clinical Epidemiology* 68, 1 (2015), 94–101. <https://doi.org/10.1016/j.jclinepi.2014.09.011>
- [99] Katarzyna Stawarz, Chris Preist, Deborah Tallon, Laura Thomas, Katrina Turner, Nicola Wiles, David Kessler, Roz Shafran, and David Coyle. 2020. Integrating the Digital and the Traditional to Deliver Therapy for Depression. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–14. <https://doi.org/10.1145/3313831.3376510>
- [100] Samantha Stonbraker, Tiffany Porras, and Rebecca Schnall. 2020. Patient preferences for visualization of longitudinal patient-reported outcomes data. *Journal of the American Medical Informatics Association* 27, 2 (2020), 212–224. <https://doi.org/10.1093/jamia/ocz189>
- [101] Richard L. Street. 2002. Gender differences in health care provider-patient communication: Are they due to style, stereotypes, or accommodation? *Patient Education and Counseling* 48, 3 (2002), 201–206. [https://doi.org/10.1016/S0738-3991\(02\)00171-4](https://doi.org/10.1016/S0738-3991(02)00171-4)
- [102] Dennis Wang, Marawin Chheang, Siyun Ji, Ryan Mohta, and Daniel A. Epstein. 2022. SnapPI: Understanding Everyday Use of Personal Informatics Data Stickers on Ephemeral Social Media. *Proceedings of the ACM on Human-Computer Interaction* 6, 2 CSCW (2022). <https://doi.org/10.1145/3555652>
- [103] Christopher H. Warner, William Bobo, Carolyn Warner, Sara Reid, and James Rachal. 2006. Antidepressant discontinuation syndrome. *American Family Physician* 74, 3 (2006), 449–456.
- [104] Peter West, Richard Giordano, Max Van Kleek, and Nigel Shadbolt. 2016. The Quantified Patient in the Doctor's Office: Challenges & Opportunities. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (CHI '16). Association for Computing Machinery, New York, NY, USA, 3066–3078. <https://doi.org/10.1145/2858036.2858445>
- [105] Peter West, Max Van Kleek, Richard Giordano, Mark J. Weal, and Nigel Shadbolt. 2018. Common Barriers to the Use of Patient-Generated Data Across Clinical Settings. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (CHI '18). Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3173574.3174058>
- [106] Naomi Yamashita, Hideaki Kuzuoka, Keiji Hirata, Takashi Kudo, Eiji Aramaki, and Kazuki Hattori. 2017. Changing Moods: How Manual Tracking by Family Caregivers Improves Caring and Family Communication. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) (CHI '17). Association for Computing Machinery, New York, NY, USA, 158–169. <https://doi.org/10.1145/3025453.3025843>
- [107] Haining Zhu, Joanna Colgan, Madhu Reddy, and Eun Kyoung Choe. 2016. Sharing patient-generated data in clinical practices: an interview study. In *AMIA Annual Symposium Proceedings*, Vol. 2016. American Medical Informatics Association, 1303.

Received July 2023; revised January 2024; accepted March 2024