JMIR Formative Research

Impact Factor (2023): 2.0
Volume 9 (2025) ISSN 2561-326X Editor in Chief: Amaryllis Mavragani, PhDc

Contents

Original Papers

Using a Proactive Telecare System to Support Independence, Health, and Well-Being in Older Adults: Feasibility Randomized Controlled Trial

Lauren Fothergill[1], PhD; Yvonne Latham[2], PhD; Niall Hayes[3], PhD; Carol Holland[1], PhD

[1]Division of Health Research, Faculty of Health and Medicine, Lancaster University, Health Innovation One, Lancaster, United Kingdom

[2]Management School, Organisation Work and Technology, Lancaster University, Lancaster, United Kingdom

[3]Leeds University Business School, Faculty of Business, University of Leeds, Leeds, United Kingdom

Corresponding Author:
Lauren Fothergill, PhD
Division of Health Research, Faculty of Health and Medicine, Lancaster University, Health Innovation One, Lancaster, United Kingdom

Abstract

Background: Proactive telecare offers services designed to reduce the occurrence of emergency situations by delivering proactive outbound calls and follow-ups and providing information and advice. By engaging regularly with users, proactive telecare may foster social connections with older adults and enable the detection of changes in needs. Telecare systems that promote active participation among older adults may also foster feelings of autonomy and self-management.

Objective: This study aimed to (1) explore the acceptability and feasibility of delivering and evaluating a proactive telecare intervention to community-dwelling older adults prior to a potential effectiveness trial and (2) evaluate the proposed eligibility criteria and estimate the potential effect size of the impact of the intervention on health and well-being outcomes to inform sample size calculations for a future trial.

Methods: An 8-week randomized pre-post feasibility study was conducted. Using a mixed methods approach, questionnaires and semistructured interviews were used to explore the feasibility and acceptability of the study. The proactive telecare system encouraged users to press an OK button once a day to confirm their well-being. If they did not respond, participants received a well-being check, and emergency contacts were notified if required. Outcomes associated with independence, health, and well-being were measured using standardized questionnaires, including health-related quality of life, mental health, and loneliness.

Results: Thirty older adults were recruited, with 13 randomized into the intervention group and 17 into the control group. The mean (SD) age of the participants was 75.4 (5.2) years; 66.7% (20/30) of the participants recruited had more than one health condition. This study achieved high retention rates (30/33, 90.9%); however, the expression of interest rate was low (52/295, 17.6%), indicating that changes to recruitment strategies are required. Effect sizes for all quantitative outcomes were small (approximately 0.2). Participants demonstrated high acceptance of the intervention, with the primary benefit cited as providing reassurance and promoting autonomy. Proactive engagement encouraged self-regulation and allowed users to control the level of support received. Those who were socially isolated reported feeling less lonely because of having additional social contact. Most participants felt the intervention would be particularly beneficial if they were experiencing poor health that significantly affected their daily activities, suggesting it may be more suited to those with limited independence. Some participants expressed anxiety about using the technology, primarily due to a lack of understanding and uncertainty in their perceived need for the device.

Conclusions: This proactive telecare system is feasible to deliver within a cohort of older adults living in the community. However, changes to recruitment approaches and implementation are needed to ensure acceptability and target numbers are achieved in a future effectiveness trial.

(JMIR Form Res 2025;9:e82152) doi:10.2196/82152

KEYWORDS

telecare; aging in place; older adults; feasibility study; acceptability

Introduction

Background

By 2030, the global population of older adults will reach 1.4 billion, which is set to increase to 2.1 billion by 2050 [1]. This demographic shift is likely to increase the number of people with health issues or disabilities, increasing the need for care of older people [2]. Promoting aging in place has become a key policy focus in the United Kingdom [3], as well as in other countries like Australia [4] and Canada [5]. This approach not only supports the health and well-being of older adults but also offers a cost-effective alternative to institutional care and helps address shortages in social care services [6]. Aging in place can be defined as "remaining living in the community, with some level of independence, rather than in residential care" [7, p. 133] or often referred to as independent living. However, aging presents physical, psychological, and social changes, such as functional decline, disability, widowhood, and increased risk of social isolation [8], which can reduce quality of life and independence [9]. Consequently, many older adults may need support to remain at home.

One approach to fostering independent living involves the utilization of telecare [10]. Telecare uses monitoring technologies, such as fall detectors and pendant alarms, to help older adults request assistance during emergencies [11]. Research has demonstrated that telecare may help promote well-being and health-related quality of life by providing a sense of security, reducing fear of falls, and increasing confidence [12,13]. However, previous research highlights concerns among older adults that telecare could be viewed as a cost-cutting strategy, potentially replacing in-person interactions [14]. This shift could contribute to greater social isolation and loneliness, factors that elevate the risk of all-cause morbidity and mortality [15]. Telecare is often implemented following incidents like falls, which can associate its use with aging and frailty [16,17]. Subsequently, telecare is frequently perceived by older adults as a final option [17], rather than a tool capable of actively promoting independence, health, and well-being.

Proactive telecare extends the support of fall detectors and pendant alarms by incorporating regular well-being monitoring [18]. This includes user-initiated check-ins via digital systems or outbound well-being calls [18]. In our previous research on proactive telecare, we found that regular engagement with older adults could facilitate early identification of emerging needs, acting as an early warning system [19]. We also found that allowing older adults to control the level of support provided through proactive telecare enhanced their perceived autonomy and improved access to social networks, particularly for those who were socially isolated [19]. Cund et al [20] similarly reported positive effects on the mental health of older adults using proactive telecare in Scotland. However, uncertainties remain about the full benefits of proactive telecare technologies to health and well-being [18].

This study focuses on a proactive telecare intervention, called *OKEachDay*, which allows users to check in daily by pressing an OK button or through outbound well-being calls. Evidence suggests that enhancing social support and increasing opportunities for social contact are effective strategies for promoting well-being and reducing loneliness [21]. This system tailors support based on the user's level of independence [19], whereby if an older adult requires additional support, it is detected when they fail to press the OK button. Conversely, if the OK button is pressed, it indicates that no further support is needed. This approach offers users control over their care [19], potentially boosting autonomy and confidence.

Given the limited research on proactive telecare [18], there were several uncertainties about conducting an effectiveness study, including intervention acceptability, compliance with daily intervention use, willingness of participants to be randomized, outcome selection, and methods for outcome collection. As a result, a feasibility study was conducted to evaluate the study's integrity for a future randomized controlled trial (RCT).

Study Objectives

The objectives of this study are to (1) assess the acceptability and feasibility of delivering a proactive telecare system to community-dwelling older adults before a full effectiveness trial (eg, randomization, assessment measures, compliance with daily intervention use); and (2) evaluate the proposed eligibility criteria and estimate the potential effect size of the impact of the intervention on health and well-being outcomes to inform a sample size calculation for a future trial.

Methods

Ethical Considerations

Ethical approval for the study was granted by the Lancaster University Faculty of Health and Medicine Ethics Committee in September 2022 (ethics board approval number: FHM-2022 - 1011-SA-1). Participants provided written consent before taking part, and data was deidentified. All participants were offered a £20 (US $22.14) shopping voucher as an appreciation for taking part in the study.

Study Design and Procedure

A mixed methods approach, incorporating both quantitative and qualitative methods, was used to effectively address the research objectives. Participants were randomized in an equal 1:1 ratio into either the intervention group, which received the proactive telecare intervention immediately for a duration of 8 weeks, or a waitlist control group, which was offered the intervention after an 8-week delay. This design aimed to address ethical concerns related to withholding a potentially beneficial intervention. The 8-week trial duration aligns with prior studies examining the feasibility of technology-based interventions aimed at supporting independence in older adults [22,23].

Data for this study were collected between 2022 and 2023 in England. Individuals interested in participating were provided with an information sheet, outlining the study details and given the opportunity to ask questions before providing written informed consent. All participants completed a baseline survey with the lead researcher (LF) a week prior to commencing the trial (Multimedia Appendix 1). The survey assessed physical health, mental health, and other outcomes related to independence. Participants were randomly allocated into the

two groups by the lead researcher using computer-generated random numbers. The same survey was administered again after participants had completed 8 weeks in the trial. Due to the nature of the intervention, neither participants nor the lead researcher could be blinded to group allocation.

Participants and Recruitment

Older adults aged 65 years and above, who lived in their own home (not a care home) and who spoke English, were invited to take part. A sample of 30 participants was aimed for to adequately estimate the effect size (potential impact of the intervention) and test the feasibility of running a larger-scale trial [24].

Participants were recruited through various channels in the Northwest of England, including local councils, an older adult research volunteer group, and local community groups. Posters were distributed to community centers, and the lead researcher presented the research study at local older adult social groups. Staff at local councils aided in recruiting participants who had been identified as at risk of loneliness and social isolation. Potential participants contacted the lead researcher (LF) if they were interested in taking part in the study. The lead researcher then arranged a time to speak with each potential participant to explain what the research involved and to assess their eligibility and capacity to participate. Eligible individuals were provided with a participant information sheet and completed a consent form before commencing the study.

Proactive Telecare Intervention

The system consisted of either a telephone or touchscreen device with an OK button for participants to press daily to confirm their well-being. Once consented to the study, participants receiving the intervention were contacted via telephone by proactive telecare staff. During this set-up call, each participant chose a preferred device (tablet or telephone), agreed on a time to press their OK button, and identified a nominated contact, often family or friends, who could be contacted if staff believed there were risks to the user. An automated reminder to press the OK button was played through the device 15 minutes before the participants' agreed time. If the participant did not press their "OK" button by the agreed cutoff time, the call center team would attempt to contact the participant to confirm their well-being, which gave an opportunity for social interaction. If staff could not reach the participant via telephone, they contacted the participant's nominated contact. In the event where nominated contacts could not be contacted, if staff believed there were critical risks to the user, emergency services were called.

Both devices have a button to press if the participant wishes to speak to the call center team, which could be used to call for help, to have a chat, or to raise other issues. Proactive telecare staff were available from 8 AM to 10 PM daily to support participants with general well-being and safety concerns. Proactive telecare staff are routinely trained in dementia awareness, suicide alertness, domestic abuse awareness, learning disability awareness, mental health awareness, and safeguarding. The service also offers additional courtesy calls to help people who may feel particularly isolated, which were offered to participants prior to taking part.

Data Collection

Quantitative Data Collection

Participant Characteristics and Intervention Use

Demographic data (age, gender, education level, ethnicity, current or previous occupation, living arrangements, current levels of care, and health conditions) and participation rates were collected. The uptake of participants on initial approach and retention of participants recruited to the study were recorded. Participants' engagement with proactive telecare was recorded to determine the feasibility of trial procedures and adherence with daily intervention use, which included the number of times a participant did not press their OK button, the number of calls between proactive telecare staff and participants, and the length of these calls.

Acceptability of Proactive Telecare

Participants in the intervention group were invited to complete an acceptability questionnaire to measure perceived usefulness, satisfaction, and ease of use of this proactive telecare using the senior technology acceptance model (STAM) 14-item scale (modified to fit the context of the intervention of interest) [25]. This measurement tool was used as it was designed to consider the needs of older adults, and it used the well-established technology acceptance model to underpin the questionnaire [26].

Health and Well-Being Outcomes

Standardized questionnaires were used to measure health-related quality of life, mental health, levels of loneliness, and perceived control and autonomy, reflecting key aspects of independence in older adults [27].

Health-related quality of life was measured using the short form-12 (SF-12) survey, due to its wide use and reliability [28]. The SF-12 measures 8 health domains, which are summarized into 2 scores, the Physical Component Summary and the Mental Component Summary. Mental well-being was measured using the Warwick-Edinburgh Mental Well-being Scale [29], which assesses hedonic well-being (happiness and life satisfaction) and eudaimonic well-being (positive psychological functioning and self-realization). It was chosen to capture positive well-being outcomes related to independence [27]. The Hospital Anxiety and Depression Scale was chosen to measure depression and anxiety [30], as it differentiates somatic symptoms that could be associated with aging as opposed to depression (eg, reduced appetite or poor sleep) [31]. The UCLA Loneliness Scale was used to measure loneliness [32], as the scale effectively measures participants' *subjective* feelings of loneliness, rather than just social isolation. Quality of life was measured using the Quality of Life Scale (CASP-19 [Control, Autonomy, Self-Realization, and Pleasure, Quality of Life Scale in Older Adults]), which assesses control, autonomy, pleasure, and self-realization in older people [33]. This tool was chosen for its relevance to an older adult population and its focus on autonomy and control, which, our previous research suggests, improve with proactive telecare use [19].

Qualitative Data Collection

Participants in the intervention group were asked to take part in a short semistructured interview upon completion of the trial. Participants from the control group who chose to use the intervention for 8 weeks after the initial waiting list period were also asked to take part in a semistructured interview. The semistructured interviews were used to explore the feasibility outcomes, including the acceptability of the proactive telecare intervention trial procedures (The interview guides can be seen in Multimedia Appendix 2). The interviews were conducted in person or over the phone if the participant preferred. All interviews were recorded with permission using an encrypted digital recorder and transcribed verbatim by the lead researcher and anonymized.

Data Analysis

Quantitative Analysis

Baseline characteristics of the intervention and control participants were summarized using descriptive statistics. Effect sizes were calculated using Hedges *g* for future use in a sample size calculation. Hedges *g* was used as it is considered to be more accurate than Cohen *d* when analyzing small sample sizes [34]. Hedges *g* was interpreted using the recommended benchmarks of 0.2 for a small effect, 0.5 for a medium effect, and 0.8 for a large effect [34]. In keeping with the aims of a feasibility study, no inferential statistics were reported.

Qualitative Analysis

Interview data were analyzed using the framework analysis method [35] to facilitate comparisons between participants and align the data with our aims. The first author led the analysis, with another researcher coding 20% of the interviews and assisting in developing the framework matrix. The analysis began with both researchers reviewing and independently coding two initial transcripts. Codes were both deductive (using concepts from STAM and the research questions) and inductive (developed from the data). They discussed the codes for relevance and meaning, leading to the development of a preliminary analytical framework. A further two transcripts were coded by both researchers using the preliminary framework, taking care to note any new themes or codes that had not been previously included. Follow-up discussions resulted in revisions to the framework to incorporate new and refined codes. The lead researcher coded the remaining transcripts, refining the framework as new codes were developed. The themes were formed using existing concepts from STAM, for example, “perceived usefulness of the intervention” and “perceived ease of use.” The final analytical framework consisted of 13 concepts, organized into four categories, each defined by a brief description (Table 1).

Table . Framework analysis for feasibility and acceptability objectives.

Concept	Description
Acceptability and usability of proactive telecare	
Perceived usefulness of intervention	The extent to which the individual feels that the technology will support their independence, improve well-being or quality of life, make them feel safer, or promote a sense of control. Additional perceived benefits may include the ability to access assistance when needed (eg, contacting designated individuals), facilitating connections to social resources, and mitigating feelings of loneliness.
Perceived ease of use of intervention	The extent to which the individual believes that the technology is easy to operate, requires little mental effort, and is clear and understandable.
Technology anxiety	Refers to an individual's hesitancy to engage with the technology, often stemming from unfamiliarity with its design, fear of making errors, or concerns about potential malfunctions.
Resistance to using technology	Refers to individuals who, despite the potential benefits of the technology, do not want to engage with it. This reluctance may be driven by financial constraints, a lack of perceived need, or reliance on alternative technologies or resources that already fulfill similar functions.
Improvements	Describes any improvements to the intervention that the participants suggest.
Appropriateness of eligibility criteria and study process	
Eligibility criteria	Describes identifying factors that highlight the appropriate people who may benefit from this technology.
Interest in taking part	Describes the participants' reasons for wanting to take part.
Acceptability of trial procedure	
Study process	Describes participants' views on the study procedures, including the randomization process, the clarity and adequacy of the information provided, and the study design.
Assessment measure	Describes participants' views on completing the surveys, suggestions for additional outcome measures that could have been included, and any support required to complete the assessments.
Compliance	Describes the daily use of proactive telecare and any issues experienced.

Results

Participant Characteristics

The mean (SD) age of participants was 75.4 (5.2) years, and all participants were identified as White British (Table 2). The majority of participants were female (76.6%, 23/30), and more than half of the participants lived alone (63%, 19/30). A small proportion of the participants had informal carers (13.3%, 4/30), and just 10% (3/30) of participants currently used other telecare devices (in this case, pendant alarms). The majority of participants (93.3%, 28/30) had at least one chronic disease or health condition.

Table . Participant descriptive characteristics.

Characteristics	Intervention group (n=13)	Control group (n=17)	Total (n=30)
Age, mean (SD)	76.7 (5.9)	74.4 (5.1)	75.4 (5.2)
Gender, n (%)			
Female	10 (76.9)	13 (76.5)	23 (76.6)
Male	3 (23.1)	4 (23.5)	7 (23.4)
Lives alone, n (%)			
Yes	9 (69.2)	10 (58.8)	19 (63.3)
No	4 (30.8)	7 (41.2)	11 (36.7)
Living arrangement, n (%)			
Private accommodation	10 (76.9)	15 (88.2)	25 (83.3)
Housing association[a]	3 (23.1)	2 (11.8)	5 (16.7)
Education, n (%)			
No qualifications	4 (30.8)	3 (17.6)	7 (23.4)
Vocational qualification	4 (30.8)	4 (23.6)	8 (26.7)
GCSE[b] or equivalent	0	2 (11.8)	2 (6.7)
A level or equivalent	1 (7.7)	3 (17.6)	4 (13.3)
Degree	3 (23)	2 (11.8)	5 (16.6)
Postgraduate	1 (7.7)	3 (17.6)	4 (13.3)
Has an informal carer, n (%)			
Yes	3 (23.1)	1 (5.9)	4 (13.3)
No	10 (76.9)	16 (94.1)	26 (86.7)
Diagnosed health condition, n (%)			
None	1 (7.7)	0	1 (3.3)
One	2 (15.4)	6 (35.3)	8 (26.7)
More than one	10 (76.9)	10 (58.8)	20 (66.7)
Prefer not to say	0	1 (5.9)	1 (3.3)
Current or previous occupation, n (%)			
Professional	7 (53.8)	5 (29.2)	12 (40)
Managerial	0	2 (11.8)	2 (6.6)
Clerical	1 (7.7)	4 (23.6)	5 (16.7)
Service and sales	4 (30.8)	1 (5.9)	5 (16.7)
Skilled agricultural	0	1 (5.9)	1 (3.3)
Trade work	1 (7.7)	4 (23.6)	5 (16.7)
Other telecare use (pendant alarm), n (%)			
Yes	2 (15.4)	1 (5.9)	3 (10)
No	11 (84.6)	16 (94.1)	27 (90)

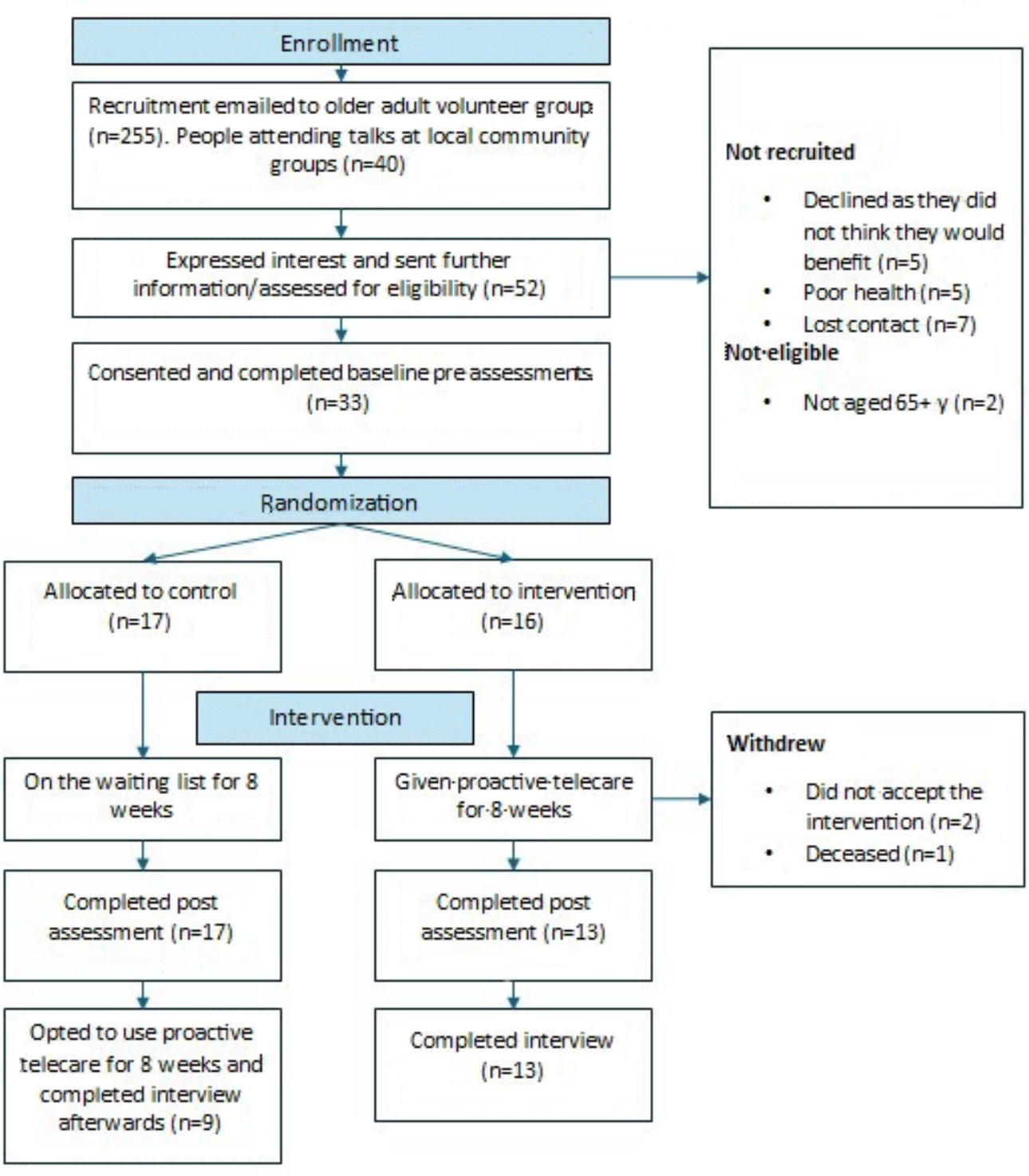
[a]Housing associations provide affordable housing options, primarily for low- and moderate-income households.

[b]GCSE: General Certificate of Secondary Education, completed at age 16 years old in the United Kingdom.

Recruitment and Retention Feasibility

Of the 295 individuals who received recruitment emails and attended an information session, 52 (17.6%) expressed initial interest. Of these, 50 (96% eligibility rate) were eligible, and 33 (66% recruitment rate) consented to participate. Reasons for nonparticipation included poor health (n=5), lack of perceived benefit from proactive telecare (n=5), and no response to follow-up emails (n=7). Among the 33 participants, 17 were randomized to the control group and 16 to the intervention group. A total of 30 participants completed the trial, yielding a 90.9% retention rate. Two intervention group participants withdrew, and one passed away during the study. Participant flow is summarized in Figure 1.

Figure 1. Consolidated Standards of Reporting Trials (CONSORT) flowchart of study recruitment, retention, and data collection.



Compliance With Daily Intervention Use

Participant engagement with the device was high, as most participants pressed their button daily or engaged with staff via phone. While almost all missed pressing their "OK" button at least once during the 8-week trial (with an average of 7 [range 1 - 49] missed presses), the level of interaction remained consistent. The average number of calls between participants and proactive telecare staff was 10 (range 3 - 57), with an average call length of 2 minutes and 27 seconds. Calls included participant-initiated contact or staff-initiated follow-ups after missed button presses. One participant did not press their OK button purposefully every day to receive a call from proactive telecare staff, as they felt socially isolated and wanted daily contact. For those requesting extra courtesy calls, the average call length was 4 minutes and 38 seconds.

Health and Well-Being Outcomes

The health and well-being outcomes are presented in Table 3. In the intervention group, self-reported physical health from the SF-12 (Physical Component Summary) improved slightly compared to the control group (unadjusted between-group difference=4.92). Both groups showed a slight reduction in self-reported mental health (Mental Component Summary). Anxiety and depression levels remained stable in both groups, and quality of life decreased slightly in both. Mental well-being improved in the intervention group, while it decreased in the control group (unadjusted between-group difference=2.54). Loneliness increased in both groups. Effect sizes for all outcomes were small (approximately 0.2). With a small effect size of 0.2 and 80% power, a sample size for an RCT would be 150 participants, based on suggestions by Faul et al [36]. There were no missing data at the two time points, as the lead researcher either read the survey to participants or checked for missing responses during data collection. Two control group participants accidentally received the intervention due to human error by the telecare company. They were kept in the control group for outcome analysis to adhere to intention-to-treat principles [37].

Table . Health and well-being outcomes at the start (pre) and after 8 weeks of the intervention (post) and control trials

	Control group (n=17), mean (SD)			Intervention group (n=13), mean (SD)			
	Pre	Post	Within-group differences, mean (SD)	Pre	Post	Within-group differences, mean (SD)	Intervention group effect size (Hedges *g*)[a]
Health-related quality of life (SF-12[b])							
PCS[c]	45.16 (10.38)	43.24 (10.69)	−1.87 (9.28)	39.15 (9.01)	42.08 (10.19)	3.05 (5.58)	0.305
MCS[d]	50.10 (11.35)	46.95 (12.25)	−3.15 (9.32)	46.92 (10.49)	43.62 (9.42)	−3.34 (6.08)	0.331
Mental well-being (WEMWBS[e])	52.76 (11.26)	51.76 (11.68)	−1 (6.72)	46.38 (11.42)	47.92 (8.87)	1.54 (7.38)	0.151
Anxiety and depression (HADS[f])	9.18 (7.34)	9.76 (7.28)	0.58 (3.89)	13.16 (7.03)	13.23 (6.76)	0.07 (3.82)	0.010
Loneliness (UCLA[g])	29.47 (11.76)	31.47 (13.34)	2.00 (9.40)	38.69 (17.51)	40.69 (16.17)	2.00 (10.90)	0.119
Quality of life (CASP-19[h])	42.71 [8.53)	41.12 [10.65)	−1.59 [7.87)	39.15 [8.41)	37.77 [7.59)	−1.38 [7.433)	0.172

[a]Hedges *g* presents the effect size for the intervention group.

[b]SF-12: short form-12.

[c]PCS: Physical Component Summary.

[d]MCS: Mental Component Summary.

[e]WEMWBS: Warwick-Edinburgh Mental Well-Being Scale.

[f]HADS: Hospital Anxiety and Depression Scale.

[g]UCLA: University of California, Los Angeles.

[h]CASP-19: Control, Autonomy, Self-Realization, and Pleasure, Quality of Life Scale in Older Adults

Acceptability of Proactive Telecare

Both the quantitative technology acceptance survey and qualitative interviews indicated a generally positive perception of proactive telecare. Participants responded to 14 items assessing acceptability, using a Likert scale (1=strongly disagree and 10=strongly agree; see Table 4). On average, participants rated proactive telecare as useful (mean 7.3, SD 3.0) and agreed that it supported their ability to live independently (mean 7.7, SD 3.0). The system was also rated as easy to use (mean 9.5,

SD 1.3). Participants generally disagreed with the statement indicating apprehension about using the technology (mean 3.1, SD 3.2). However, participants on average agreed that the cost of the intervention was a concern (mean 6.7, SD 3.7).

Table . Technology acceptance survey responses. Participants were asked to indicate their agreement with the technology acceptance statements below, using a Likert scale (1=strongly disagree, 10=strongly agree).

Technology acceptance statements	Scores, mean (SD)
Attitudinal beliefs	
Using proactive telecare enhanced your ability to live independently	7.7 (3.0)
You found proactive telecare useful in your daily activities	7.3 (3.0)
You like the idea of using proactive telecare	8.0 (2.4)
Control beliefs	
Proactive telecare was easy to use	9.5 (1.3)
You could complete a task using proactive telecare if there was someone to demonstrate how	9.1 (2.2)
Your financial status does not limit your activities in using proactive telecare	6.7 (3.7)
When you want or need to use proactive telecare, it is accessible to you	9.4 (1.5)
Technology anxiety	
You feel apprehensive about using proactive telecare	3.1 (3.2)
You hesitate to use proactive telecare for fear of making mistakes you cannot correct	2.7 (2.8)
Health conditions	
How are your general health conditions? (with 1 being very poor and 10 being very good)	6.7 (2.4)
How well are you able to concentrate? (with 1 being very uneasy and 10 being very easy)	7.8 (1.5)
How satisfied are you with your personal relationships? (with 1 being very unsatisfied and 10 being very satisfied)	8.3 (1.9)
How satisfied are you with the support received from friends and family? (with 1 being very unsatisfied and 10 being very satisfied)	8.4 (2.2)
How satisfied are you with your quality of life? (with 1 being very unsatisfied and 10 being very satisfied)	8.1 (1.5)

The findings from the quantitative acceptance survey were corroborated with the qualitative findings. Three themes were interpreted from the data: (1) perceived usefulness of proactive telecare, (2) perceived ease of use, and (3) technological anxiety and resistance.

Perceived Usefulness of Proactive Telecare

Participants indicated that the most valuable aspect of the proactive telecare system was the reassurance provided by having a remote support network monitoring their physical and mental well-being, along with the ability to request help if needed.

> *I would describe it really as a comfort blanket, you just know that it's as though somebody's looking out for you and I think that's a nice feeling when you're getting older, just that you don't want to be alone.* [Participant 18]

Participants echoed the importance of proactive engagement in providing reassurance of safety. Participants also emphasized the benefit of being proactive in promoting self-initiation and self-regulation.

> *It's reassurance, isn't it, I think it's a psychological trigger. I think it's a good thing, I really do.* [Participant 12]

One participant highlighted the value of the flexibility of the intervention because users had choice and control over the level of support provided if they missed their OK button, in comparison to a pendant alarm where activating it indicates an emergency in an "all or nothing" approach to support.

> *you press that [pendant alarm] for help, that's like saying it's an emergency, do I really need it? Just to say like in the morning yeah, I'm OK today, that's better I think.* [Participant 30]

While some participants felt they were slightly too young to require a proactive telecare system, several expressed surprise at the perceived benefits of having a remote monitoring team overseeing their well-being. They noted that, in the event of an incident such as a fall where the "OK" button was not pressed,

the system would initiate a response to check on their well-being, which many participants found reassuring. However, a few were skeptical about the system's effectiveness in promoting in-home safety and preferred alternatives like pendant alarms or mobile phones for requesting help. Approximately half of the participants suggested using a pendant alarm in conjunction with the proactive telecare system to enhance the ability to request help when needed.

> *I do think that people who are in danger of falling need a falls alarm as well.* [Participant 3]

Participants characterized the proactive telecare staff as friendly, empathetic, and supportive, contributing to a sense of being cared for and emotionally reassured.

> *It was nice. It felt to me as if they really cared about me, it felt personal, I could feel as if that lady or that young man was ringing me because they were concerned about me.* [Participant 4]

Two participants reported experiencing feelings of loneliness prior to their involvement in the study and subsequently chose to receive additional courtesy calls from the proactive telecare staff. These participants described forming positive interpersonal relationships with the staff and emphasized that the human contact provided through these conversations was more meaningful to them than the reassurance of safety alone.

> *The people at [proactive telecare service] are beautiful people who are lovely, I think it has helped and, like I say, them ringing me twice a week, it's really been nice. I'll miss it really; you don't feel as lonely.* [Participant 10]

Although most participants declined the additional courtesy calls, as they believed their existing levels of social interaction were sufficient, participants recognized the potential value of such calls in offering social support to older adults experiencing isolation.

Perceived Ease of Use

All participants reported that the proactive telecare intervention was easy to set up and use. Most opted to use the touchscreen device, while two participants preferred the telephone-based version. The telephone devices were installed in person by the proactive telecare staff, whereas the touchscreen devices were delivered by post, with setup instructions provided remotely via telephone.

> *I'm a technophobe, I'm useless with things like that, but no it didn't bother me at all. It was simple to use. I plugged it in the dining room and just did it every morning in the allotted time and it was just very, very, very simple to use. (User talking about using the tablet)* [Participant 9]

Most participants reported that the device was not intrusive and was not burdensome, which was viewed as positive and facilitated the development of a routine for pressing the OK button. However, some participants stated that the requirement to engage with proactive telecare daily was cumbersome and became tedious, particularly when they forgot to press the button and subsequently received follow-up calls.

> *I actually feel quite relieved, I haven't got to do it anymore [after the trial]. So perhaps I felt, it did tie me down - that I've got to remember to do it.* [Participant 5]

Most participants reported occasionally forgetting to press their button but found the automated reminders helpful as a gentle prompt. They also appreciated the flexibility of being able to choose a time that suited their daily schedule, including the option to press the button up to 6 hours before the scheduled time.

> *it didn't matter if I did sleep a bit longer, if I didn't wake up till nine o'clock I could still press it and it was alright. (User's cut off time was 10 AM).* [Participant 21]

This flexibility supported participants in remembering to press the OK button and enhanced the system's accessibility by accommodating diverse daily routines.

Technological Anxiety and Resistance

Although most participants found the technology easy to use, some reported initial apprehension when first engaging with proactive telecare, expressing fear of pressing incorrect buttons and making mistakes. One participant chose the telephone version of the device because they found it more familiar.

> *I'm not good with a tablet, I thought at least with the telephone I know there were them three things and that's all I needed to press.* [Participant 11]

Some participants stated that their lack of understanding of how the technology worked enhanced their anxiety about making a mistake and that more comprehensive explanations would have been beneficial. For some, the unfamiliar design of the tablet contributed to feelings of confusion and apprehension.

> *In the early days, I touched it in the wrong place to try and bring the screen back up again. And because I wasn't familiar with the screen, I touched the alert call. And then I couldn't see in my panic, how to cancel it. And, you know, felt really quite stupid.* [Participant 5]

Some participants stated that they would have preferred a face-to-face explanation of the technology, noting that they learned more effectively through visual demonstrations. However, most were satisfied with the telephone-based introduction to proactive telecare.

Some participants reported that they would only consider adopting proactive telecare after experiencing functional decline, a decision influenced both by the cost of the device and by associations between the technology and aging or declining health.

> *I'm only eighty and I can still get about, but somebody who couldn't get out of the house or needed help, it would be ideal for them.* [Participant 18]

Some participants felt that a perceived need for technological support was essential for engaging with the intervention. The presence of chronic conditions or disabilities was also seen as

a key factor that could make individuals more likely to benefit from proactive telecare.

> *I think it would be very handy for those who aren't quite, very well. There is knowing that there's a backup there if anything starts to go wrong.* [Participant 17]

In contrast, some participants viewed the proactive telecare system as a potential precursor to using a pendant alarm, appreciating that it did not require constant bodily wear and that they did not yet feel "ready" to adopt a pendant.

One participant noted a desire to continue using the intervention but indicated that financial constraints, due to the cost of other telecare services, prevented them from affording both devices.

> *I'd have liked to have kept it, you know, but now I've got this to pay for this pendant, it's too expensive to have both.* [Participant 13]

Appropriateness and Acceptability of Trial Procedures

Eligibility Criteria

The inclusion criteria for this study were intentionally broad, based on the assumption that older adults aged over 65 years have diverse physical and emotional needs. When asked about their motivations for participating, most individuals expressed a general interest in contributing to research and a desire to give back to their local community. Additionally, some participants were motivated by curiosity about telecare technologies and a wish to explore potential personal benefits.

> *I thought I'd like to test a system where I could make contact if I did inadvertently fall or, in any way become unsafe at home, and it came up.* [Participant 2]

Participants supported the use of broad eligibility criteria, viewing the decision to adopt the technology as dependent on individual perceived need. They identified several circumstances in which proactive telecare could be beneficial, including for individuals who are housebound, living with chronic conditions or disabilities, have a history of falls, or experience limited social support. While some participants emphasized the benefits of the system for those living alone, others highlighted its value for individuals cohabiting with others, particularly if both people had chronic conditions.

> *I think it fits with us quite well because we've both got problems and you don't know if we're going to finish up in hospital and then the other's on their own all of a sudden, it's an insurance.* [Participant 13]

Randomization and Assessment Measures

Most participants expressed satisfaction with being randomized into either the intervention or control group. However, a small number of participants preferred to be in the intervention group, so they could use the technology straight away. One participant, who reported feeling extremely isolated, expressed a strong desire to begin the intervention immediately.

Many participants reported that the trial questionnaires were acceptable and easy to complete. Nevertheless, a few participants described difficulty in answering some of the questions due to the subjectivity of some questions. One participant, for example, described challenges in answering questions that required recalling emotions or feelings experienced over the past few weeks.

> *they asked you to remember the last week or the last four weeks. And at my age, you don't remember the last week or the last four weeks very clearly.* [Participant 8]

Participants reported that the length of the survey was not burdensome, and the questions were deemed relevant to the study subject.

Discussion

Key Findings

This study assessed the acceptability and feasibility of evaluating a proactive telecare intervention in older adults living in the community. The trial sustained low dropout rates and successful collection of outcome variables. However, initial expression of interest in the study was low. Our mixed methods study suggested that proactive telecare was generally acceptable to participants; however, some participants indicated hesitancy as to whether this intervention was beneficial to them, so adaptations to the recruitment process should be explored. The trial procedures, including randomization and completing questionnaires, were feasible and acceptable to participants. Nevertheless, two instances of randomization contamination occurred, suggesting that revised procedures may be needed in a future full-scale trial.

Due to uncertainty about who would benefit from proactive telecare, the study adopted broad eligibility criteria: living in the community and being over 65 years old. This resulted in a high eligibility rate (96%), but only 17.6% of those contacted expressed interest. The low engagement may reflect the broad recruitment strategy, which targeted older adults generally who may not have perceived a need for telecare. Previous studies suggest that some older adults associate telecare with frailty, which may deter uptake [38,39]. Participants considered the criteria appropriate, emphasizing the importance of individual choice in assessing suitability. However, participants identified groups who may particularly benefit, including those who are housebound, have limited mobility, have chronic conditions, have disabilities, or lack social support. Future trials may be more effective if they target these specific populations.

Most participants found the study processes feasible and acceptable, with a high retention rate of 90.9%, consistent with findings from other studies involving older adults [40,41]. The high retention rate could be due to low participant burden imposed by the study. However, two participants withdrew due to issues with the technology. One participant was unable to establish a regular routine of use, and another was concerned about the risk of falling when standing to press the button, a concern also observed in other feasibility research involving frailer older adults at risk of falls trialing new technology [42]. Previous research suggests that telecare should integrate into individual contexts, routines, and abilities [43,44]. Providing

additional support to assess users' needs and preferences prior to implementation may help improve retention.

Many participants reported positive experiences of the study, including recruitment, randomization, and data collection methods. Many cited altruism as their reasons for taking part, which has been seen in other research involving older people [45]. While participants were generally comfortable with being randomized, a few expressed a preference for being placed in the intervention group rather than the control group. Two instances of contamination occurred, where participants allocated to the control group were provided with the intervention due to human error by the proactive telecare staff. Despite this, the participants remained in the control group to adhere to the intention-to-treat analysis [46]. Similar contamination issues have been noted in research evaluating interventions in primary care and community settings, where preventing contamination in RCT designs can be challenging [47]. In a full-scale trial, cluster randomization could be used, where groups of older adults are randomized instead of individual participants. This approach, within a defined setting such as in an assisted living environment where people live independently with additional support, could help overcome this issue.

Many participants found the outcome measures appropriate and the questionnaires easy to complete, though some noted the questions were subjective and at times difficult to answer. These findings indicate a need to optimize the measurement tools by simplifying and clarifying questionnaire items. Completing questionnaires with a researcher was viewed as helpful for clarifying meanings, suggesting researcher support may benefit future trials. Previous studies have also highlighted the importance of increased personal contact by researchers to support older adults taking part in research [45].

Only descriptive statistics were calculated to assess the feasibility of the RCT plan, as recommended for feasibility studies [48]. Standard deviations and effect sizes were calculated to be used in a future sample size estimation, as suggested by previous research [49]. Descriptive analyses found small improvements in mental well-being in the intervention group compared to the control group, which are similar to results that have been noted in other studies on telecare interventions [13]. These changes may reflect increased perceptions of safety and security, which in turn may improve perceived health and mental well-being [13]. In both groups, quality of life decreased, and loneliness scores increased; this may have been influenced by completing the surveys, which could increase participants' awareness of loneliness or their quality of life. A larger-scale trial would be required to better understand the effects of proactive telecare on health and well-being.

This study found that proactive telecare was both acceptable and feasible for older adults. Participant engagement with the system was high, as all participants pressed their OK button daily or engaged with staff via phone. The most frequently reported benefit was a sense of reassurance, which has been noted in other studies on telecare [50]. While some participants did not feel they currently needed telecare due to their perceived independence, many still valued the reassurance that someone was checking in. Proactive engagement acted as a psychological prompt, encouraging self-regulation and allowing users to control the level of support received, as noted in our previous qualitative research on proactive telecare [19]. In contrast, existing literature on other telecare interventions, particularly monitoring technologies such as ambient sensors, emphasizes older adults' concerns regarding privacy when using telecare and monitoring technologies [38,51]. The Farr Point report suggested that proactive well-being checks may be accepted by older people who are resistant to using other telecare devices due to associated stigma [18] or concerns about reduced perceptions of control and privacy [52]. This highlights the potential for proactive telecare interventions to promote autonomy and self-management, but further evaluation is required to fully understand this.

Most participants reported that the intervention was easy to set up and use and that it would be appropriate for older people who may not have experience in using similar technologies due to the simplicity of the system. There were a few participants who felt some apprehension toward the technology and would have benefited from having an in-person demonstration, which should be considered in a future trial. Wu et al [53] suggest that older adults often report a lack of knowledge of technology, which can result in apprehension. Product demonstrations are suggested to enable participants to trial and test out devices to gain further knowledge and confidence about the usability and usefulness of technologies.

For participants experiencing social isolation, the opportunity for social interaction was the most valued component of proactive telecare. Brief courtesy calls provided a sense of connection, with the supportive approach of staff helping to ease feelings of loneliness. However, the calls in this study lasted only around 5 minutes. In contrast, more intensive proactive telecare models, such as that examined by Cund et al [20], which involved longer, regular well-being calls, have been associated with improved mental well-being. A future RCT is warranted to evaluate the effectiveness of social support provided by proactive telecare on loneliness. Further research should also explore the optimal duration and nature of contact needed to reduce loneliness and promote well-being.

The findings from this study highlight key considerations for the future evaluation of proactive telecare. While participants valued the reassurance provided, those at higher risk may still require reactive devices for emergency support. This underscores the need for a flexible, person-centered approach, as a single technological solution is unlikely to meet the diverse needs of the older population. Future evaluations should consider how proactive telecare can be effectively integrated within local health and social care systems, including the use of other digital interventions like other telecare devices, telehealth, and telemedicine. This approach aligns with regional digital strategies in the United Kingdom, which aim to provide more preventive, proactive, and fully integrated telecare services [54]. Future evaluations should also consider how to offer proactive telecare beyond the research period, as some participants may have benefitted from continued use.

Strengths and Limitations

A key strength of this study was its mixed methods design, which provided both quantitative outcomes and qualitative insights into participants' experiences with proactive telecare. Standardized measurement tools were utilized to measure independent outcomes, which can be compared across the literature. However, several limitations should be noted. The sample lacked diversity, as all participants were White British, indicating a need for more inclusive recruitment in future trials. The 8-week duration may have been too short to observe full benefits, though longer follow-ups may affect retention. Additionally, the study focused on one type of proactive telecare, limiting generalizability.

Conclusions

This study highlights key considerations for designing a future RCT of a proactive telecare system for older adults living in the community. The intervention was generally well-received, offering reassurance and, for socially isolated individuals, a sense of social connection. Study procedures were feasible and acceptable, though improvements in recruitment and implementation procedures are suggested to maximize uptake. The data from this study have provided valuable considerations for refining and justifying the design of a future effectiveness trial.

Acknowledgments

We would like to thank all participants who took part in this study. The views expressed in this paper are those of the authors and not necessarily those of the National Health Service or Department of Health and Social Care (National Institute for Health and Care Research).

Funding

This study was funded by the National Institute for Health and Care Research (NIHR) Applied Research Collaboration North West Coast.

Data Availability

The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

Authors' Contributions

The lead author, LF, conducted this research as part of her PhD in Health Research at Lancaster University. LF was supervised by CH, NH, and YL. YL coded 20% of the transcripts and assisted with the analysis. All supervisors contributed to the concept and design of the study, reviewed final analysis and interpretation of data, critically commented on the manuscript, and agreed to the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Quantitative survey questions
[DOCX File, 31 KB - formative_v9i1e82152_app1.docx]

Multimedia Appendix 2
Qualitative interview guide
[DOCX File, 17 KB - formative_v9i1e82152_app2.docx]

References

1. WHO. Ageing and health. 2024. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2025-06-22]
2. Xi JY, Liang BH, Zhang WJ, et al. Effects of population aging on quality of life and disease burden: a population-based study. Glob Health Res Policy 2025 Jan 14;10(1):2. [doi: 10.1186/s41256-024-00393-8] [Medline: 39810282]
3. Sixsmith A, Sixsmith J. Ageing in place in the United Kingdom. Ageing Int 2008 Sep;32(3):219-235. [doi: 10.1007/s12126-008-9019-y]
4. O'Loughlin K, Kendig H, Browning C. Challenges and opportunities for an ageing Australia. In: Ageing in Australia 2017, Vol. 16:1-10. [doi: 10.1007/978-1-4939-6466-6_1]
5. Thinking about aging in place. Government of Canada. 2025. URL: https://www.canada.ca/en/employment-social-development/corporate/seniors-forum-federal-provincial-territorial/aging.html [accessed 2025-06-22]

6. Peek STM, Luijkx KG, Rijnaard MD, et al. Older adults' reasons for using technology while aging in place. Gerontology 2016;62(2):226-237. [doi: 10.1159/000430949] [Medline: 26044243]
7. Davey J, de JV, Nana G, Arcus M. Accommodation options for older people in Aotearoa/New Zealand. : Centre for Housing Research Aotearoa/New Zealand; 2004 URL: https://thehub.sia.govt.nz/assets/documents/accommodation_options_for_older_people_in_aotearoa_new_zealand.pdf [accessed 2025-12-05]
8. Chen K. Use of gerontechnology to assist older adults to cope with the COVID-19 pandemic. J Am Med Dir Assoc 2020 Jul;21(7):983-984. [doi: 10.1016/j.jamda.2020.05.021] [Medline: 32546395]
9. Alnajjar F, Khalid S, Vogan AA, Shimoda S, Nouchi R, Kawashima R. Emerging cognitive intervention technologies to meet the needs of an aging population: a systematic review. Front Aging Neurosci 2019;11:291. [doi: 10.3389/fnagi.2019.00291] [Medline: 31798439]
10. Goodwin N. The state of telehealth and telecare in the UK: prospects for integrated care. J Integrated Care 2010 Nov 26;18(6):3-10. [doi: 10.5042/jic.2010.0646]
11. Gathercole R, Bradley R, Harper E, et al. Assistive technology and telecare to maintain independent living at home for people with dementia: the ATTILA RCT. Health Technol Assess 2021 Mar;25(19):1-156. [doi: 10.3310/hta25190] [Medline: 33755548]
12. Stewart LSP, McKinstry B. Fear of falling and the use of telecare by older people. Br J Occup Ther 2012 Jul;75(7):304-312. [doi: 10.4276/030802212X13418284515758]
13. Hirani SP, Beynon M, Cartwright M, et al. The effect of telecare on the quality of life and psychological well-being of elderly recipients of social care over a 12-month period: the whole systems demonstrator cluster randomised trial. Age Ageing 2014 May;43(3):334-341. [doi: 10.1093/ageing/aft185] [Medline: 24333802]
14. Mort M, Roberts C, Pols J, Domenech M, Moser I, EFORTT investigators. Ethical implications of home telecare for older people: a framework derived from a multisited participative study. Health Expect 2015 Jun;18(3):438-449. [doi: 10.1111/hex.12109] [Medline: 23914810]
15. Fakoya OA, McCorry NK, Donnelly M. Loneliness and social isolation interventions for older adults: a scoping review of reviews. BMC Public Health 2020 Feb 14;20(1):129. [doi: 10.1186/s12889-020-8251-6] [Medline: 32054474]
16. Astell AJ, McGrath C, Dove E. 'That's for old so and so's!': does identity influence older adults' technology adoption decisions? Ageing Soc 2020 Jul;40(7):1550-1576. [doi: 10.1017/S0144686X19000230]
17. Bentley CL, Powell LA, Orrell A, Mountain GA. Making telecare desirable rather than a last resort. Ageing Soc 2018 May;38(5):926-953. [doi: 10.1017/S0144686X16001355]
18. Proactive telecare services: final study report. Digital Health & Care Scotland. 2020. URL: https://www.tec.scot/wp-content/uploads/Evaluating-Proactive-Telecare-Final-v2-23.01.23.pdf [accessed 2025-12-05]
19. Fothergill L, Holland C, Latham Y, Hayes N. Understanding the value of a proactive telecare system in supporting older adults' independence at home: qualitative interview study among key interest groups. J Med Internet Res 2023 Dec 14;25(1):e47997. [doi: 10.2196/47997] [Medline: 38096023]
20. Cund A, Hendry A, Murray A, et al. Evaluating the introduction of proactive telecare in Scotland. Int J Integr Care ;22(S3):150. [doi: 10.5334/ijic.ICIC22067]
21. Masi CM, Chen HY, Hawkley LC, Cacioppo JT. A meta-analysis of interventions to reduce loneliness. Pers Soc Psychol Rev 2011 Aug;15(3):219-266. [doi: 10.1177/1088868310377394] [Medline: 20716644]
22. Kramer LL, Mulder BC, van Velsen L, de Vet E. Use and effect of web-based embodied conversational agents for improving eating behavior and decreasing loneliness among community-dwelling older adults: protocol for a randomized controlled trial. JMIR Res Protoc 2021 Jan 6;10(1):e22186. [doi: 10.2196/22186] [Medline: 33404513]
23. Sattar S, Haase K, Penz K, et al. Feasibility of a remotely delivered strength and balance training program for older adults with cancer. Curr Oncol 2021 Nov 2;28(6):4408-4419. [doi: 10.3390/curroncol28060374] [Medline: 34898562]
24. Lancaster GA, Dodd S, Williamson PR. Design and analysis of pilot studies: recommendations for good practice. J Eval Clin Pract 2004 May;10(2):307-312. [doi: 10.1111/j..2002.384.doc.x] [Medline: 15189396]
25. Chen K, Lou VWQ. Measuring senior technology acceptance: development of a brief, 14-item scale. Innov Aging 2020;4(3):igaa016. [doi: 10.1093/geroni/igaa016] [Medline: 32617418]
26. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
27. Hillcoat-Nallétamby S. The meaning of "independence" for older people in different residential settings. J Gerontol B Psychol Sci Soc Sci 2014 May;69(3):419-430. [doi: 10.1093/geronb/gbu008] [Medline: 24578371]
28. Huo T, Guo Y, Shenkman E, Muller K. Assessing the reliability of the short form 12 (SF-12) health survey in adults with mental health conditions: a report from the wellness incentive and navigation (WIN) study. Health Qual Life Outcomes 2018 Feb 13;16(1):34. [doi: 10.1186/s12955-018-0858-2] [Medline: 29439718]
29. Tennant R, Hiller L, Fishwick R, et al. The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. Health Qual Life Outcomes 2007 Nov 27;5:63. [doi: 10.1186/1477-7525-5-63] [Medline: 18042300]
30. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983 Jun;67(6):361-370. [doi: 10.1111/j.1600-0447.1983.tb09716.x] [Medline: 6880820]

31. Grover S, Sahoo S, Chakrabarti S, Avasthi A. Anxiety and somatic symptoms among elderly patients with depression. Asian J Psychiatr 2019 Mar;41:66-72. [doi: 10.1016/j.ajp.2018.07.009] [Medline: 30054249]
32. Russell DW. UCLA Loneliness Scale (Version 3): reliability, validity, and factor structure. J Pers Assess 1996 Feb;66(1):20-40. [doi: 10.1207/s15327752jpa6601_2] [Medline: 8576833]
33. Hyde M, Wiggins RD, Higgs P, Blane DB. A measure of quality of life in early old age: the theory, development and properties of a needs satisfaction model (CASP-19). Aging Ment Health 2003 May;7(3):186-194. [doi: 10.1080/1360786031000101157] [Medline: 12775399]
34. Hedges LV. Distribution theory for Glass's estimator of effect size and related estimators. J Educ Stat 1981 Jun;6(2):107-128. [doi: 10.3102/10769986006002107]
35. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013 Sep 18;13:117. [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
36. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/bf03193146] [Medline: 17695343]
37. Gupta SK. Intention-to-treat concept: a review. Perspect Clin Res 2011 Jul;2(3):109-112. [doi: 10.4103/2229-3485.83221] [Medline: 21897887]
38. Karlsen C, Ludvigsen MS, Moe CE, Haraldstad K, Thygesen E. Experiences of community-dwelling older adults with the use of telecare in home care services: a qualitative systematic review. JBI Database System Rev Implement Rep 2017 Dec;15(12):2913-2980. [doi: 10.11124/JBISRIR-2017-003345] [Medline: 29219874]
39. Chaudhuri S, Kneale L, Le T, et al. Older adults' perceptions of fall detection devices. J Appl Gerontol 2017 Aug;36(8):915-930. [doi: 10.1177/0733464815591211] [Medline: 26112030]
40. Zhang Y, Leuk JSP, Teo WP. Domains, feasibility, effectiveness, cost, and acceptability of telehealth in aging care: scoping review of systematic reviews. JMIR Aging 2023 Apr 18;6:e40460. [doi: 10.2196/40460] [Medline: 37071459]
41. Choi YK, Thompson HJ, Demiris G. Use of an internet-of-things smart home system for healthy aging in older adults in residential settings: pilot feasibility study. JMIR Aging 2020 Nov 10;3(2):e21964. [doi: 10.2196/21964] [Medline: 33170128]
42. Tarraf RC, Suter E, Arain M, et al. Using integrated technology to create quality care for older adults: a feasibility study. Inform Health Soc Care 2019 Sep;44(3):246-261. [doi: 10.1080/17538157.2018.1496090] [Medline: 30102117]
43. Mortenson WB, Sixsmith A, Beringer R. No place like home? Surveillance and what home means in old age. Can J Aging 2016 Mar;35(1):103-114. [doi: 10.1017/S0714980815000549] [Medline: 26742899]
44. Fothergill L, Hayes N, Latham Y, Hamilton J, Ahmed S, Holland C. Understanding how, for whom and under what circumstances telecare can support independence in community-dwelling older adults: a realist review. BMC Geriatr 2025 Jan 27;25(1):59. [doi: 10.1186/s12877-024-05650-6] [Medline: 39871177]
45. Hurst C, Dismore L, Granic A, et al. Feasibility of engaging older adults living with multiple long-term conditions, frailty, and a recent deterioration in health in a study of lifestyle: protocol for the LiLL-OPM study. J Frailty Sarcopenia Falls 2023 Jun;8(2):127-135. [doi: 10.22540/JFSF-08-127] [Medline: 37275663]
46. Pitkala KH, Strandberg TE. Clinical trials in older people. Age Ageing 2022 May 1;51(5):afab282. [doi: 10.1093/ageing/afab282] [Medline: 35526845]
47. Smith SM, Soubhi H, Fortin M, Hudon C, O'Dowd T. Interventions for improving outcomes in patients with multimorbidity in primary care and community settings. Cochrane Database Syst Rev 2012 Apr 18;2016(4):CD006560. [doi: 10.1002/14651858.CD006560.pub2] [Medline: 22513941]
48. Tickle-Degnen L. Nuts and bolts of conducting feasibility studies. Am J Occup Ther 2013;67(2):171-176. [doi: 10.5014/ajot.2013.006270] [Medline: 23433271]
49. Suresh K, Chandrashekara S. Sample size estimation and power analysis for clinical research studies. J Hum Reprod Sci 2012 Jan;5(1):7-13. [doi: 10.4103/0974-1208.97779] [Medline: 22870008]
50. Johannessen TB, Storm M, Holm AL. Safety for older adults using telecare: perceptions of homecare professionals. Nurs Open 2019 Jul;6(3):1254-1261. [doi: 10.1002/nop2.328] [Medline: 31367452]
51. Birchley G, Huxtable R, Murtagh M, Ter Meulen R, Flach P, Gooberman-Hill R. Smart homes, private homes? An empirical study of technology researchers' perceptions of ethical issues in developing smart-home health technologies. BMC Med Ethics 2017 Apr 4;18(1):23. [doi: 10.1186/s12910-017-0183-z] [Medline: 28376811]
52. Berridge C. Active subjects of passive monitoring: responses to a passive monitoring system in low-income independent living. Ageing Soc 2017 Mar;37(3):537-560. [doi: 10.1017/S0144686X15001269] [Medline: 28239211]
53. Wu YH, Damnée S, Kerhervé H, Ware C, Rigaud AS. Bridging the digital divide in older adults: a study from an initiative to inform older adults about new technologies. Clin Interv Aging 2015;10:193-200. [doi: 10.2147/CIA.S72399] [Medline: 25624752]
54. Going digital: a technology enabled care strategy for 2025. : Lancashire County Council; 2020 URL: https://council.lancashire.gov.uk/documents/s200313/Appendix%20A.pdf [accessed 2025-12-06]

Abbreviations

CASP-19: Control, Autonomy, Self-Realization, and Pleasure, Quality of Life Scale in Older Adults
RCT: randomized controlled trial
STAM: senior technology acceptance model

Edited by A Schwartz, M Balcarras; submitted 11.08.25; peer-reviewed by C Asuzu, M Lennon; revised version received 05.11.25; accepted 25.11.25; published 31.12.25.

Please cite as:
Fothergill L, Latham Y, Hayes N, Holland C
Using a Proactive Telecare System to Support Independence, Health, and Well-Being in Older Adults: Feasibility Randomized Controlled Trial
JMIR Form Res 2025;9:e82152
URL: https://formative.jmir.org/2025/1/e82152
doi:10.2196/82152

Reach, Engagement, and Acceptability of a Subclinical Telehealth Service for Spanish-Speaking Adults: Retrospective Mixed Methods Pilot Study

Marvyn R Arévalo Avalos[1], PhD; Julio Fu[2], MBA; Adrian Aguilera[1,3], PhD; Luis Suarez[2], MBA

[1]School of Social Welfare, University of California, Berkeley, 120 Haviland Hall, Berkeley, CA, United States

[2]Sanarai, Chicago, IL, United States

[3]Department of Psychiatry and Behavioral Sciences, University of California, San Francisco, San Francisco, CA, United States

Corresponding Author:
Marvyn R Arévalo Avalos, PhD
School of Social Welfare, University of California, Berkeley, 120 Haviland Hall, Berkeley, CA, United States

Abstract

Background: There is a gap in mental health care among Latino/x and Spanish-speaking communities, and the care that is available is often difficult to access, lacks cultural nuance, and results in low engagement and satisfaction.

Objective: The study aimed to evaluate the reach, adoption, and acceptability of digital Spanish-language psychosocial and emotional wellness services among Latino/x adults offered by the digital health company Sanarai.

Methods: Data included in this study were collected between August 2020 and September 2024 by Sanarai as part of its ongoing services. Quantitative data sources included individual customers' appointment data, individual session payment data, and customer satisfaction data. Qualitative data were obtained from transcribed notes of telephone or video-based user interviews conducted by Sanarai staff between August 2020 and May 2024.

Results: Between August 2020 and September 2024, Sanarai served 6163 users (n=3662, 59.42% women participants) across all 50 US states, with the highest concentration of participants in Texas and California. Results showed 94% (n=5793) of users scheduled a first appointment within 1 week, with 43% (n=2650) doing so within 1 day. Over 62.60% (n=3858) of participants engaged in two or more sessions, attending an average of 8.94 (SD 13) sessions over 110 days (SD 169). The platform delivered a total of 36,858 appointments, including individual and couples sessions. Only 22.47% (n=1385) of users responded to a customer satisfaction survey for a total of 2287 distinct responses; among this subgroup, session satisfaction was high with an average satisfaction rating of 4.88 out of 5.0 (SD 0.49) and a Net Promoter Score of +85. Nearly all responses (n=2174, 95.06%) expressed intent to schedule another session, but these results should be interpreted with caution, given the low response rate. Qualitative interviews with 30 users (n=21, 70% women) revealed a diverse user base. Many users reported prior mental health service experiences, while one-third were new to care. Participants cited cost, cultural fit, language access, and convenience as key reasons for choosing Sanarai over local services. Users highlighted the platform's affordability, scheduling flexibility, and provider professionalism as central to their positive experiences.

Conclusions: These findings underscore the value of culturally responsive, accessible online mental health care for Spanish-speaking communities.

(JMIR Form Res 2025;9:e80026) doi:10.2196/80026

KEYWORDS

digital health; Latino/x; mixed methods; Spanish-language mental health care; telehealth services

Introduction

Latino/x individuals, who represent 20% of the US population [1], experienced a high level of unmet need for mental health services [2]. Estimates suggested that nearly 1 in 2 (41% - 49%) of Latino/x people with a perceived need for mental health services also experienced unmet mental health care [3-5]. Notably, 58% of foreign-born Latino/x adults and 12% of the US-born adults preferred a Spanish-speaking health care provider or a Latino/x health care provider (47% among foreign-born and 20% among the US-born Latino/x) [6]. However, according to the American Psychological Association, only 5% of the psychologist workforce offered Spanish-language services [7], and only 5% to 6% of the mental health counselor and psychologist workforce identified as Latino/x [7,8]. There was a critical need for services and resources that addressed disparities in mental health care among Latino/x adults in the United States [9-11]. Cost, stigma, and language access were often cited as barriers to care. The cost of mental health services

in the United States poses a barrier for uninsured and underinsured communities [11], which is characterized by many Spanish-speaking Latino/x individuals. Stigma was also cited as a possible factor deterring individuals from seeking care due to not wanting to feel ostracized, identified as mentally ill, or perceptions of mental health needs as weakness [12,13]. The large unmet mental health needs of vulnerable populations [4,5,14] are detrimental to individuals and families. There was an urgent need to develop and evaluate solutions that were accessible and culturally relevant for Latino/x and Spanish-speaking communities.

The rapidly growing field of digital mental health, particularly online therapy or digital mental health coaching, has the potential to meet the needs of vulnerable populations. Digital solutions also tended to be more private and could help reduce stigma. However, many of these digital efforts were limited in reaching and engaging Spanish-speaking and Latino/x populations. Among Latino/x adults, acceptability and engagement with digital mental health interventions were often low [15,16], thus negatively impacting the potential benefits of digital mental health solutions for this population. Reasons for low engagement and satisfaction were that most digital health solutions were not linguistically or culturally congruent, particularly for Spanish-speaking adults, and were difficult for Latino/x adults to access (eg, cost, time, and availability) [11,17]. The COVID-19 pandemic gave rise to the rapid expansion and adoption of telehealth mental health services across the United States [18], but these gains were not observed across all demographic groups. For example, mental health treatment centers in communities with high proportions of Latino/x residents were twice as *unlikely* to adopt telehealth services relative to communities with smaller proportions of Latino/x residents [19]. These discrepancies in the adoption of telehealth services have contributed to greater disparities in unmet mental health needs among Latino/x adults, particularly those who preferred Spanish-language services. In addition, a national evaluation of digital health technologies among Spanish speakers in the United States found that digital literacy, limited availability and appropriateness of information in Spanish, and the lack of Spanish-speaking interpreters limited patients' adoption of online health portals and telehealth services; whereas, facilitators of digital technology adoption included access to interpreters and culturally and linguistically tailored Spanish-language materials and health care providers [20]. Collectively, these results indicated that the lack of linguistically and culturally appropriate services and providers was a significant barrier to accessing care among Spanish-speaking communities.

One way to address the shortage of Spanish-speaking providers in the United States was to connect with mental health providers in predominantly Spanish-speaking countries via digital health technologies. For example, private sector digital health companies, such as Sanarai, connected Spanish-speaking clients in the United States to licensed mental health professionals in Latin America via telehealth with the goal of providing linguistically and culturally competent subclinical mental health services [21,22]. These types of programs using providers from Latin America for mental health support have not been formally evaluated, but there was evidence to suggest that these providers could fill existing gaps in care. For instance, evidence from primary care programs, such as the Licensed Physicians from Mexico Pilot Program (LPMPP) implemented in California in 2002, suggested that working with providers abroad could help address existing domestic gaps. The LPMPP was approved with the goal of taking 30 licensed physicians from Mexico to provide medical care to underserved Latino/x and Spanish-speaking communities at nonprofit community health centers throughout the state to address shortages of health care providers and to increase linguistically and culturally competent care [23]. An evaluation of the LPMPP indicated that the program met the intended goals of improving access to linguistically and culturally competent care and resulted in high acceptability of the LPMPP physicians among patients, particularly among monolingual Spanish-speaking patients [23]. The results of this program were promising and provided some evidence of the benefits of addressing gaps in care and the shortage of Spanish-speaking mental health professionals with providers from Spanish-speaking countries.

As noted earlier, private sector companies were bridging gaps in access to linguistically and culturally grounded mental health care for Latinos/x and Spanish-speaking communities. Although private companies were often not set up to test specific hypotheses, it was important to apply implementation frameworks to understand if interventions were acceptable, feasible, and addressing needs. Acceptability and engagement were two important and interrelated implementation constructs that were associated with the effectiveness of digital mental health interventions [24-27]. Results from existing research indicate that acceptability impacted engagement, and vice versa, and that both constructs were influenced by the user, intervention, and technology factors [24,27]. For example, facilitators include user beliefs toward technology and mental health, perceived fit and usefulness of the intervention, ease of integration of the digital solutions into one's life (eg, cost and availability), and privacy and confidentiality of the digital tools [24,26,28,29]. Much of the recent extant digital mental health research focused on the use of smartphone apps or other technology-delivered, asynchronous, self-paced interventions and was limited in its inclusion of Spanish-speaking populations [25,28,30,31]. Thus, there was a need to evaluate whether telehealth-delivered mental health supportive services for Spanish-speaking adults were acceptable, engaging, and potentially viable digital health solutions to address gaps in care for this vulnerable population.

The purpose of this mixed methods retrospective pilot study was to evaluate and describe the implementation of a telehealth-delivered, Spanish-language, psychosocial and emotional support service among Spanish-speaking adults. The mixed methods approach was useful in health services research [32] as the quantitative data helped assess the magnitude of impact or outcomes (eg, reach and engagement) while the qualitative data helped understand the process and content of use. The inclusion of rich interview data provided a more comprehensive understanding of the multidimensional facets of intervention acceptability and engagement, which might not otherwise be obtained via survey data. This study examined the

reach, engagement, and acceptability of the Spanish-language telehealth services by analyzing quantitative and qualitative data gathered via the routine services provided by the digital health company Sanarai. In addition, the results of this study aimed to provide evidence regarding the feasibility of leveraging mental health providers from Spanish-speaking countries to meet existing gaps in care in linguistically and culturally appropriate mental health services for Latino/x and Spanish-speaking adults.

Methods

Sanarai

Sanarai is a digital health company that arranges for the delivery of psychosocial and emotional supportive services to adults (aged ≥18 y) via telehealth technologies from licensed Spanish-speaking professionals from Latin America to Spanish-speaking adults located in the United States and other countries. These services are primarily provided direct-to-consumers and target Latino/x and Spanish-speaking populations residing in the United States with the goal of increasing access to high-quality, culturally congruent, and affordable subclinical mental health supportive services. The company services include: advice related to general mental health and wellness; life coaching and recommendations for setting personal goals; emotional support related to interpersonal relationships, sexuality, and gender; and emotional support related to feelings of anxiousness, stress, and depression. Sanarai addresses the subclinical population, emphasizing support for individuals dealing with challenges that may not meet the clinical threshold but still significantly impact their well-being. The focus of the services is to provide early intervention and preventive support. The company was founded in 2020.

Sanarai has developed a simple workflow to onboard users into its platform. First, users can browse the up-to-date provider database found on the Sanarai website to select the provider of their choice (step 1) and choose their desired appointment time (step 2). Then, users provide basic contact and payment information (step 3) to book the appointment, which is confirmed via email and conducted via Zoom (Zoom Communications, Inc; step 4). Users have the flexibility of changing their provider at any time in the process.

Sanarai users may sign up for an introductory 30-minute initial consultation or a full 50-minute session, either as an individual or as a couple. All sessions are conducted via live video calls on the video platform, Zoom. At the time of this evaluation, the initial consultation was priced at US $20.00 and served as an opportunity for the user to get to know their provider and their working style. The individual and couples' sessions, US $49.00 and US $59.00, respectively, are focused on addressing the psychosocial and emotional wellness needs of the user. Sanarai's providers are licensed mental health professionals from Latin America with at least a master's degree and a minimum of 5 years of clinical experience. However, Sanarai notifies its users that despite the providers' educational background and clinical expertise, Sanarai's services are not considered psychiatric, psychological, or mental health therapy services as traditionally provided by a licensed mental health provider in the United States. Thus, Sanarai providers do not diagnose, prescribe medication, or accept health insurance. As needed, Sanarai will provide resources and referrals to users who need a higher level of care or are seeking services related to mental health diagnoses or medication management.

Data Sources and Analysis

Sanarai has routinely collected quantitative and qualitative data as part of its ongoing services. Data included in this study were collected between August 2020 and September 2024. Quantitative data sources included three sources: (1) individual customers' appointment data, (2) individual sessions payment data, and (3) customer satisfaction data. Specifically, all appointment and user data (1) were sourced directly from the booking platform, Acuity Scheduling, via automatic export. This system was consistently used throughout the entire evaluation period, ensuring uniformity. This appointment data included the date and time of each appointment, the type of appointment, the price and amount paid for each appointment, and whether the user was referred via an organizational partnership. Users' first name was used to generate a gender variable consistent with Spanish-language name norms (ie, first names ending with the vowels "a" or "e" were coded as "Woman participants," those ending with "o" were coded as "Man participants," and a manual review was done to double check the coding). Payment data (2) was used to extract the user's zip code, if available. These quantitative data served as the basis to analyze reach and adoption by examining trends in users' engagement and session data. In addition, after each individual session, Sanarai automatically emailed users inviting them to participate in a brief web-based satisfaction survey (3) using a standardized Google Form, which included questions about satisfaction with services, willingness to recommend the services, and the Net Promoter Score measure. These data provided a snapshot of the acceptability of the service.

Quantitative data analysis was conducted using R software (R Foundation for Statistical Computing) [33] and included computing descriptive statistics for the number of users, the number and types of sessions attended, session engagement metrics, and the geographic distribution of users across the continental United States. Though this study was focused on a description of implementation outcomes and not hypothesis testing, exploratory analysis was conducted to examine usage patterns by gender and by time to schedule the first appointment.

Qualitative data were obtained from transcribed notes of telephone or video-based user interviews conducted by Sanarai staff (3 members of the executive team) between August 2020 and May 2024. The primary goal of these interviews was internal product improvement, not formal research, yet these interviews were structured and used a questionnaire based on the "Jobs To Be Done" framework. The "Jobs To Be Done" framework focused on understanding the motivations and decisions that drive a user to select a product or service to accomplish their goals or meet their needs (aka meet the requirements of the job they want done) [34]. The interviews focused on understanding the user experiences, motivations, and preferences for seeking mental health services with Sanarai. Participants were recruited via generalized outreach to the entire user base on a quarterly

basis. Participation in these interviews was based on 100% self-selection, and everyone who expressed interest was interviewed. At times, the company prioritized interviewing "super users" (high session count) or users who had recently stopped seeking services in order to understand specific use patterns. Users who opted in to participate in these interviews were offered a discount of up to 15% on their next session. These qualitative data served as the basis to analyze acceptability.

We examined the user interview transcript and conducted reflexive thematic analysis (TA) [35,36] to examine users' attitudes and experiences with seeking and accessing mental health services, including services provided by Sanarai. The reflexive TA was conducted by the first author, who began by reviewing the transcripts and independently coding the data. Codes were generated based on an inductive data-driven and semantically focused approach [36] with the goal of summarizing patterns in the data via themes that best represented the beliefs and experiences of the participants. A key concept in reflexive TA was that themes emerge at the intersection of the data and the researcher's subjectivity and research skills [36]. Thus, to provide an extra layer of validation of the qualitative findings, the codes, themes, and narrative presentation of the qualitative results were reviewed and approved by the second and fourth authors, who had more familiarity with the interview process and qualitative interviews.

Ethical Considerations

This study was reviewed by the University of California, Berkeley Institutional Review Board (IRB #2024-08-17735) and granted exempt status as it satisfies the federal and UC Berkeley requirements under category 70: research that involves no greater than minimal risk to subjects. The data used in this study were collected by Sanarai, as part of enrollment in its services, users provide broad consent to the potential use of their data for program evaluations and secondary analysis. Data were not fully anonymized, as the data sources included a unique Sanarai ID comprised of the email address the participant used to enroll in the services. To safeguard participant information, data were shared directly between Sanarai and the first author and stored under university-sponsored password-protected accounts. Data were accessed under a data transfer agreement between the University of California, Berkeley and Sanarai. Participants who completed the user interviews were offered a discount of up to 15% for their next session. No additional participant compensation was provided. The study's conceptualization, analysis, and interpretations were conducted independently by the first author.

Results

Reach

During the 4-year period between August 2020 and September 2024, Sanarai provided psychosocial and emotional wellness support services to a total of 6163 users. The majority of users were women (n=3662, 59.4%), and the remaining were men (n=2501, 40.6%). In terms of geographic region, payment data indicated that 65.7% (n=4049) of users made payments using a credit card obtained in the United States; 6.27% (n=387) made payments using a credit card obtained outside of the United States. For the remaining 28.02% (n=1727) of users, their geographic region was unknown as the payment data did not include country or zip code information. Within the US-based sample, 97.06% (n=3930) of the users had available zip code data, which were used to estimate participant locations within the United States (Figure 1). In sum, all 50 states of the United States, the District of Columbia, and Puerto Rico were represented in the sample, yet 55% of the users were concentrated in just 5 states: Texas (n=621, 15.34%), California (n=567, 14%), Florida (n=539, 13.31%), New York (n=246, 6.08%), and New Jersey (n=190, 4.69%). All other states represented 43.64% (n=1767) of the users, and 2.94% (n=119) were unknown. Sanarai users were reached via online advertisements, partnerships with community-based organizations, and word of mouth (recommendations by family and friends). During this evaluation period, Sanarai's user base and number of active users grew an average of 47% quarter over quarter (Figure 2).

Figure 1. Participant locations in the continental United States. Each blue dot represented an approximate location for each participant. Location estimates were based on available participant zip code data.

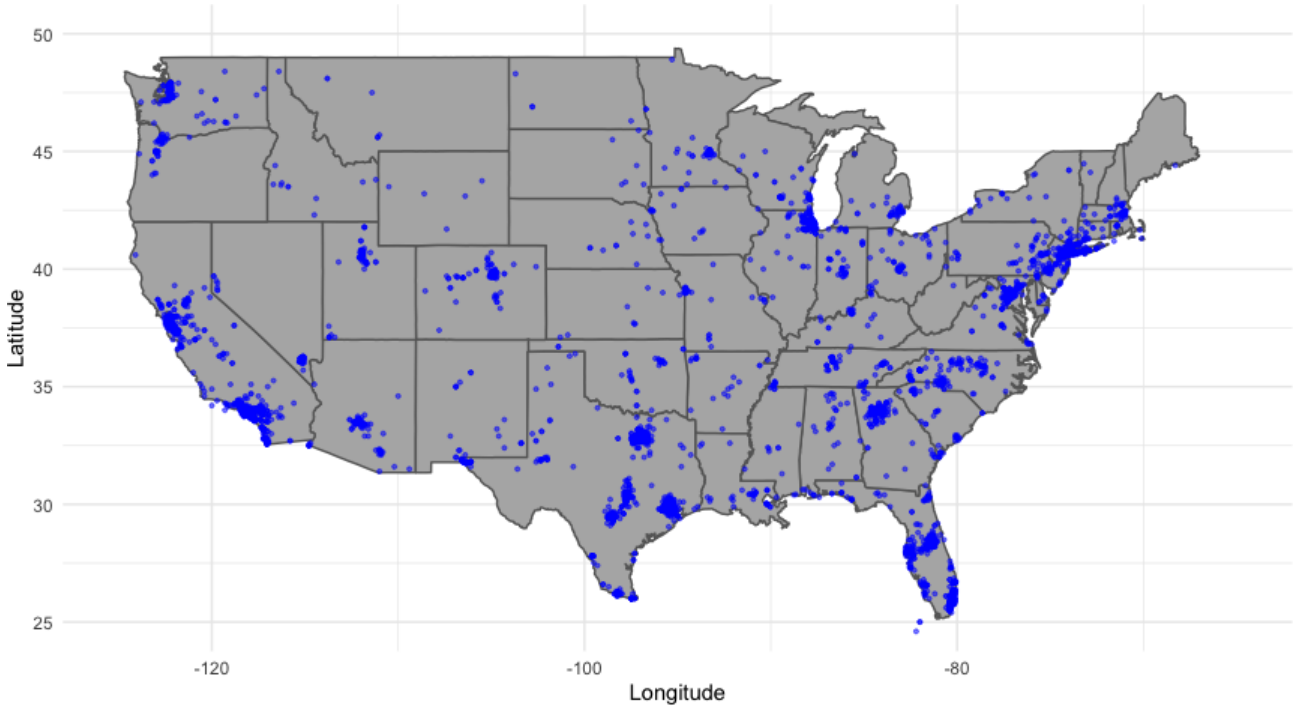
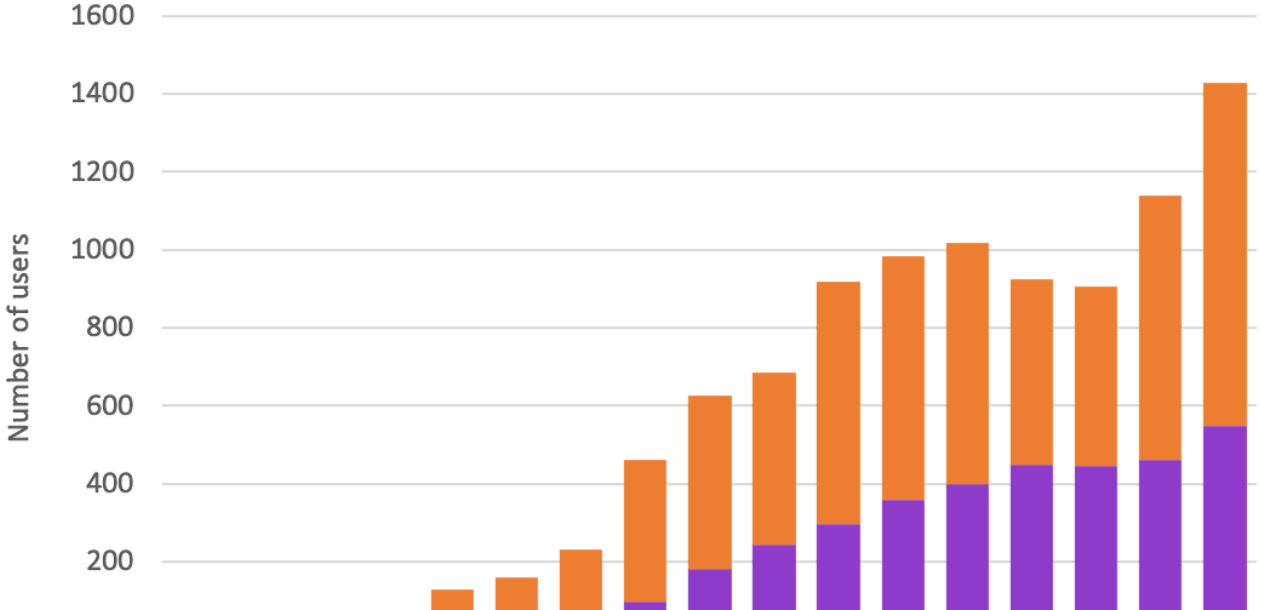


Figure 2. Quarterly trends of Sanarai users. This stacked bar graph represents the total numbers of active users in the online telehealth service, Sanarai, for each calendar quarter starting from quarter 3: July 2020-September 2020 (20 Q3) through quarter 3: July 2024-September 2024 (24 Q3). The orange (top) bars corresponded to new users in the service, the purple (bottom) bars corresponded to existing users. The table below the bar graph includes the counts for each new and existing user by quarter.



	20 Q3	20 Q4	21 Q1	21 Q2	21 Q3	21 Q4	22 Q1	22 Q2	22 Q3	22 Q4	23 Q1	23 Q2	23 Q3	23 Q4	24 Q1	24 Q2	24 Q3
New users	9	31	42	48	108	122	194	366	445	441	622	627	618	476	460	676	878
Existing users	0	4	12	21	20	38	39	97	181	245	297	358	399	449	446	462	550

Engagement

Among the 6163 users served by Sanarai, the first point of contact for 53.33% (n=3287) was an initial consultation, and the rest (n=2876, 46.67%) began with a full session. In terms of appointment availability, 43% (n=2650) of users scheduled their first appointment within 1 day or less, and an additional 30% (n=1849) scheduled within 2 to 3 days. Overall, the vast majority of users (n=5793, 94%) scheduled a first appointment

within 7 days, and a minority scheduled their first appointment after more than 7 days (n=370, 6%).

Regarding engagement pattern and service use types, 36.26% (n=2235) of users attended 1 appointment only (either a consultation or 1 single full session), and 1.14% (n=70) had consultations without a single full session; whereas 62.6% (n=3858) of users engaged with Sanarai services at least two or more times via a combination of consultation and sessions or sessions alone (Table 1). Among this subsample of users with greater than or equal to 2 sessions of any type, we calculated the total number of sessions attended, days between sessions, and time between first and last session. Full descriptive statistics for these engagement metrics are reported in Table 2. In sum, users attended an average of 8.94 sessions, with an average of 12.4 days between sessions, and remained engaged for an average of 110 days. Finally, during the evaluation period, Sanarai provided a total of 36,858 appointments—including individual consultations (n=3563, 9.67%), individual sessions (n=31 905, 86.6%), couples sessions (n=1179, 3.2%), and couples initial consultations (n=211, 0.57%). This represented approximately 768 appointments per month.

Table . Engagement of Sanarai users.

Engagement	Value, n (%)
First appointment type	
Initial consultation	3104 (50.37)
Couples initial consultation	183 (2.97)
Individual session	2643 (42.88)
Couples session	233 (3.78)
Service use types	
Consultations only	1214 (19.70)
Single session	1091 (17.70)
Multiple sessions	1700 (27.58)
Consultation + sessions	2158 (35.02)

Table . Engagement metrics.

	Median (IQR)	Mean (SD)	Minimum	Maximum	SE (95% CI)
Total number of sessions	5 (3-9)	8.94 (13.02)	2	236	0.21 (8.53-9.35)
Days between sessions	7 (4.50-11.50)	12.40 (20.96)	0	521	0.34 (11.74-13.06)
Days engaged	40 (14-121)	110.49 (168.95)	0	1445	2.72 (105.15-115.82)
Time to schedule first appointment (any type)	2 (1-4)	2.80 (2.62)	1	27	0.04 (2.72-2.88)

To further contextualize engagement patterns, we conducted two exploratory analyses. First, we conducted an independent samples 2-tailed *t* test to examine whether the total number of sessions and the total number of days engaged differed by gender. The results indicated that there were no statistically significant differences in total number of sessions between women participants (mean 8.99, SD 12.75) and men participants (mean 8.86, SD 13.41; $t_{3338.6}$=0.29; 95% CI −0.72 to 0.97; *P*=.77). Similarly, we did not observe any statistically significant differences in the number of days engaged (women participants: mean 114.27, SD 169.89; men participants: mean 105.15, SD 167.57; $t_{3476.6}$=1.66; 95% CI –1.67 to 19.92; *P*=.10). Finally, we conducted a correlation between time to schedule first appointment and total number of sessions to examine if prompt services were associated with greater intervention engagement and found a very small and weak negative correlation (*r*=−0.038, 95% CI −0.063 to −0.013).

Acceptability

Overview

Sanarai's session satisfaction survey was completed by 22.47% (n=1385) of its users, resulting in 2287 distinct responses. These data showed a Net Promoter Score of +85, which indicates that a large portion of respondents were likely to recommend the services to friends and family. In addition, there was an average session satisfaction rating of 4.88 (SD 0.49; rating 1=very bad and 5=very good), and 95.06% (n=2174) of responses indicated "Yes" to scheduling a new session. These data provided a snapshot of high acceptability for Sanarai services among this user base, but this should be interpreted with caution given the low response rate.

Qualitative data (user interview data) were analyzed from 30 participants. These data were collected between August 3, 2020, and May 29, 2024. There was an overrepresentation of women participants (n=21, 70%) versus men participants (n=9, 30%). Participants had an average age of 30.54 (SD 6.81; range

19 - 46) years. Participants were diverse in terms of country of origin, with Mexico representing 43.33% (n=13) of the sample. The US state of residence distribution was similar to the full sample (Table 3), and the length of time residing in their state ranged from 2 months to 30 years (mean 6.73, SD 6.60 y). One-third of users (n=10, 33.33%) reported not having any past mental health service experiences. The remaining users (n=20, 66.67%) reported having prior experience with mental health services, which varied from long-term exposure (n=8), recent exposure only (n=8), or exposure in one's home country (n=5).

Table . Demographics of user interviews.

Characteristic	Value, n (%)
Gender	
Woman	21 (70)
Man	9 (30)
Country of origin	
Mexico	13 (43.33)
United States	4 (13.33)
Venezuela	4 (13.33)
Dominican Republic	3 (10)
Other[a]	6 (20)
State residing	
California	5 (16.67)
New York	5 (16.67)
Florida	3 (10)
Texas	3 (10)
Other[b]	14 (46.67)

[a]Ecuador, Guatemala, Colombia, Puerto Rico, and Spain.

[b]Georgia, North Carolina, Connecticut, Illinois, Massachusetts, Michigan, Minnesota, New Jersey, New Mexico, Pennsylvania, Tennessee, and Washington State.

Reflexive TA Themes

Reasons for Seeking Services

Users reported a wide range of reasons for which they were seeking mental health support. These reasons included a history of mental health concerns, general mental health stressors, transitions related to stress, family-related stress, and self-growth. Several users reported seeking support to manage symptoms related to depression, anxiety, binge eating, insomnia, or other undisclosed history of mental health issues. On the other hand, some users were seeking support to address general mental health stressors, increased emotionality and vulnerability, grief, or acute life stressors. In terms of transition-related stress, users identified immigration-related stress, missing family or their home country, coping with the COVID-19 pandemic, and loneliness as reasons for needing support. Family stressors included marital and relationship conflicts, concerns about supporting children's mental health, and concerns about one's family. Finally, a minority of users reported seeking services as a form of personal improvement and self-growth.

Use of Local Resources

In discussing participants' decision to seek online services, users described their experiences and beliefs about seeking local, in-person resources to meet their needs. Participants discussed four main themes that presented as barriers to using local resources, which included cost, accessibility, cultural fit, and modality fit. First, most participants reported that local mental health services were too expensive; that mental health services were not covered under their health insurance; or that even if those services were included, the co-pays and deductibles would be comparable to Sanarai's service pricing and, thus, considered cost a significant barrier to seeking mental health services in their local communities.

Accessibility was another prominent issue reported by participants, which included barriers with transportation to in-person services, difficulty navigating the US health care system, lack of available providers, and long wait times to receive services. A third barrier to using local resources was related to the cultural fit of services, with users reporting that despite looking for help, they were unable to find Spanish-speaking language providers or culturally competent care. Finally, users reported issues with a lack of modality fit, meaning that when they sought services for mental health, they were referred to psychiatry, which focused too much on medication prescription when they preferred talk therapy, or that even when providers advertised having available in-person appointments, they would only accept patients for video-based appointments.

Value of Sanarai or Virtual Mental Health Support

Participants reported that Sanarai and online mental health support service value was driven by its ability to overcome barriers related to accessing local services. Specifically, many users endorsed the ability to receive Spanish-language services and culturally congruent care from a trustworthy and professional source as a main driver for seeking online services via Sanarai. Accessibility was another important consideration for users, as they stated they liked the ease of scheduling sessions, flexibility in scheduling, ease of making payments, and choice about selecting or switching their provider.

Cost was also cited as a reason for choosing Sanarai versus other resources. Although to a lesser extent than the previous themes, some users reported that Sanarai's services were affordable and that receiving a free initial consult motivated them to engage in services. For example, some users noted issues related to affordability and health insurance. Finally, a minority of users endorsed a variety of other topics, including finding the website attractive and professional, having reliable technology to conduct sessions, and appreciating the privacy and confidentiality of Sanarai's services, specifically this service not being connected to a user's medical record.

Experience With Sanarai Providers

Users reported choosing their specific provider because their online profile fit the patient's needs. For example, some users identified a provider who met their demographic or identity preferences (eg, ethnic background, gender, and religion). Others thought the providers' years of experience and areas of expertise or focus were a good fit to address their concerns. Users also reported feeling comfortable communicating with their providers due to their level of professionalism, empathy, helpfulness, and understanding expressed throughout sessions. However, a minority of users reported having mixed reactions toward their provider and reported feeling as if their provider was not as attentive or providing the right type of support. These users acknowledged that when they were not feeling comfortable with their provider, they knew they were able to switch providers, yet some chose to terminate services instead.

Discussion

Principal Findings

This study evaluated the reach, engagement, and acceptability of a telehealth Spanish-language psychosocial and emotional wellness service offered by a digital health company, Sanarai. The service delivery model focused on connecting users to Spanish-speaking licensed mental health providers from Mexico and Argentina to receive subclinical mental health supportive services via telehealth. The results of this mixed methods evaluation suggest that Sanarai services are accessible and acceptable as evidenced by the reach of the services to over 6000 users across the United States, an average duration of 9 sessions attended per user, and some preliminary indication of high satisfaction rates. Factors that may contribute to Sanarai's model acceptability include providing prompt (next-day session availability), relatively affordable rates compared to market rates in the United States (US $49-$59 for individual and couples' sessions), offering culturally and linguistically congruent mental health services for Latino/x adults, and having an accessible and easy-to-use technology platform.

A central goal of the Sanarai model of mental health support was to increase the availability of language-concordant mental health providers for Spanish-speaking adults. The geographic reach of Sanarai's services is representative of the widespread need for professional mental health care services in the United States. According to the Health Resources and Services Administration, there is a significant shortage of mental health professionals at the national level, as only 26.4% of the mental health professional workforce needs were met as of March 31, 2025 [37]. According to the report, in the top 5 US states where Sanarai had the highest user enrollment, the percent of needs met were 15.6% (New York), 22.4% (California), 23.7% (Florida), 31.3% (Texas), and 52.7% (New Jersey) [37], indicating an ongoing significant need for additional mental health professionals. Further, between 2014 and 2019, there was a 17.8% decrease in facilities that offered mental health treatment in Spanish despite a 4.5% growth in the Latino/x population during the same time period [38]. The importance of having access to Spanish-speaking health care providers cannot be overstated, as current estimates suggest only 5% to 6% of providers identify as Latino/x and were able to provide Spanish-language services [7,8], and many Latino/x people preferred to receive health care services in Spanish, particularly when discussing emotional concerns [6,39,40].

In terms of availability, Sanarai users waited on average 2.87 days to meet with their provider. The availability of Sanarai's providers is drastically different from the availability of psychologists in the United States. According to a 2023 national survey by the American Psychological Association, more than half of psychologists (56%) reported having no openings for new patients or having a waitlist of 3 months or more [41]. The risk of long wait times to receive services may result in a patient's worsening mental health [42] and greater health care expenditures, or patients may lose interest in engaging in services despite having an ongoing need. In this study, we did not find a significant association between appointment waiting time and total engagement with services. Sanarai users in this study attended an average of 9 sessions. Comparatively, other studies have shown that Latino/x individuals attend an average of 5 sessions for remote telephone-based services [43] and 8 sessions for in-person mental health care [15]. Finally, women typically endorse more positive health-seeking attitudes for mental health issues than men [44], are more likely to report needing and accessing mental health services [45], and are more likely to engage with digital mental health interventions [24]; however, in this study, we did not find any statistically significant differences in session engagement or duration of treatment by gender, suggesting that both men and women accessed and engaged with Sanarai at similar rates. Examining additional demographic variables, such as age and educational background, can further help explain patterns in service use among diverse segments of the population.

The qualitative data provides a rich description of the various reasons why Sanarai users engaged with the services and helps contextualize the quantitative findings. Contrary to the

quantitative data in which there was a more balanced breakdown of gender (60% women participants and 40% men participants), the respondents of the qualitative surveys were predominantly women (70%); thus, these results should be interpreted with caution due to the potential risk of bias and limited generalization. The respondents' reasons for seeking services were varied in intensity and ranged from managing symptoms related to mental health concerns to addressing situational stressors and self-growth. Severity of symptoms has been associated with greater engagement with digital mental health interventions [9,24] and may partly explain the engagement rates observed in this study. Users noted a wide range of barriers to accessing in-person mental health services within their communities. Many of the issues raised were also well documented in the extant literature, including barriers due to language fluency, cultural differences between patient and provider, lack of access to low-cost mental health services, and perceived lower quality of mental health care [5,45,46]. For the participants in this study, these barriers may have been the precursor to engagement and acceptability of digital mental health services.

Acceptability of digital mental health services in general, and Sanarai's services specifically, can be understood within the multifaceted framework of health care intervention acceptability [26,27]. Specifically, users endorsed ease of access, scheduling, technology navigation, and availability of providers (eg, low burden of engagement) as a significant facilitator to engaging with Sanarai. Similarly, the users perceived the services to be affordable and a good fit given their cultural values, goals, and language preference, and perceived the providers as trustworthy and helpful (ie, concurrent acceptability), which are all facilitators associated with increasing acceptability, engagement, and effectiveness of digital mental health interventions [24,26,27,39,40]. These findings provided evidence for the value of leveraging mental health providers from the Spanish-speaking countries to deliver mental health supportive services to Spanish-speaking adults in the United States. In addition, it is unclear how the acceptability of the service would impact clinical outcomes; thus, future research may include an outcomes-based evaluation or a randomized clinical trial design.

This study had several limitations. First, the data were primarily descriptive, based on users from one private telehealth company, and did not have a comparison group. These limitations limit generalizability, as it was possible that the sample included in this research was not representative of the broader Spanish-speaking populations. Also, due to the limited availability of demographic data collected via the appointment data, we were unable to examine the representativeness of the sample and how known factors such as income or health insurance status impacted users' participation with Sanarai services. Similarly, at the time of this evaluation, Sanarai was beginning to collect clinical effectiveness data; thus, we were not able to assess the clinical impact of these online services on users' mental health outcomes and how these outcomes are related to engagement and acceptability. This limitation remained an area of priority for future work. Regarding acceptability, though the user satisfaction survey was shared regularly with all users active in the service—less than 25% of the total sample participated in this survey—potentially limiting the generalizability of these findings to the entire sample and inflating positive findings as a result of self-selection bias. In addition, the results focused on acceptability were based on the qualitative data analysis, which included an overrepresentation of women relative to men and thus did not generalize. Similarly, the user interviews were conducted by company staff, and despite using a standardized protocol, it was possible that users responded with bias. Areas for future directions included collecting clinical outcomes data; increasing the collection of demographic data to facilitate greater exploration of reach, engagement, and acceptability patterns; increasing representativeness of samples; and increasing response rates to satisfaction surveys.

Conclusions

Overall, Sanarai's model appears to be meeting a need for mental health services for Spanish-speaking Latino/x communities by delivering accessible, prompt, and linguistically and culturally congruent mental health supportive services.

Acknowledgments

The authors would like to acknowledge Vanessa Villaverde, Dave Chokshi, David Stark, and Richard Mockler for their support and feedback throughout the course of this research. Generative artificial intelligence tools were not used in the writing or analysis of this manuscript.

Funding

This work was supported by a grant from the California Health Care Foundation (#G-33613).

Authors' Contributions

Conceptualization: MRAA
Data curation: MRAA, JF
Formal analysis: MRAA
Funding acquisition: MRAA, JF, LS
Investigation: JF, LS
Project administration: MRAA

Validation: JF, LS
Writing – original draft: MRAA
Writing – review & editing: MRAA, JF, AA, LS

Conflicts of Interest

MRAA and AA have provided independent consulting to digital health companies. JF and LS are employees of Sanarai.

References

1. Piña G, Martinez G. Key facts about US Latinos. Pew Research Center. 2025 Oct 22. URL: https://www.pewresearch.org/short-reads/2025/10/22/key-facts-about-us-latinos/ [accessed 2025-11-07]
2. Cook BL, Trinh NH, Li Z, Hou SSY, Progovac AM. Trends in racial-ethnic disparities in access to mental health care, 2004–2012. Psychiatr Serv 2017 Jan 1;68(1):9-16. [doi: 10.1176/appi.ps.201500453] [Medline: 27476805]
3. Padilla-Frausto DI, Babey SH, Wolstein J, Tse HW. Uncovering unique challenges: variation in unmet mental health needs among Latinx ethnic groups in California. UCLA Center for Health Policy Research. 2021 Jul. URL: https://healthpolicy.ucla.edu/sites/default/files/2023-04/unmetmentalhealthneeds-latinxethnicgroups-policybrief-jul2021.pdf [accessed 2025-11-07]
4. Loeb TB, Viducich I, Smith-Clapham AM, et al. Unmet need for mental health services utilization among under-resourced Black and Latinx adults. Fam Syst Health 2023 Jun;41(2):149-159. [doi: 10.1037/fsh0000750] [Medline: 36521110]
5. Garcia ME, Hinton L, Gregorich SE, Livaudais-Toman J, Kaplan C, Karliner L. Unmet mental health need among Chinese and Latino primary care patients: intersection of ethnicity, gender, and English proficiency. J Gen Intern Med 2020 Apr;35(4):1245-1251. [doi: 10.1007/s11606-019-05483-9] [Medline: 31667737]
6. Funk C, Lopez MH. Hispanic Americans' trust in and engagement with science. Pew Research Center. 2022 Jun 14. URL: https://www.pewresearch.org/science/2022/06/14/hispanic-americans-trust-in-and-engagement-with-science/ [accessed 2025-11-07]
7. 2021 survey of health service psychologists: technical report. : American Psychological Association; 2022 Oct URL: https://www.apa.org/workforce/publications/health-service-psychologists-survey/full-technical-report.pdf [accessed 2025-12-18]
8. Table A-11 employed persons by detailed occupation, sex, race, and Hispanic or Latino ethnicity current population survey, annual average 2024. US Bureau of Labor Statistics. 2024. URL: https://www.bls.gov/cps/cpsaat11.htm [accessed 2025-11-07]
9. Bucay-Harari L, Page KR, Krawczyk N, Robles YP, Castillo-Salgado C. Mental health needs of an emerging Latino community. J Behav Health Serv Res 2020 Jul;47(3):388-398. [doi: 10.1007/s11414-020-09688-3] [Medline: 32002728]
10. Wells K, Klap R, Koike A, Sherbourne C. Ethnic disparities in unmet need for alcoholism, drug abuse, and mental health care. Am J Psychiatry 2001 Dec;158(12):2027-2032. [doi: 10.1176/appi.ajp.158.12.2027] [Medline: 11729020]
11. Mental health disparities: Hispanics and Latinos. American Psychiatric Association. 2017. URL: https://www.psychiatry.org/getmedia/5d5b46c3-9b28-4e50-9b73-7c738e0901d8/Mental-Health-Facts-for-Hispanic-Latino.pdf [accessed 2024-05-20]
12. Benuto LT, Gonzalez F, Reinosa-Segovia F, Duckworth M. Mental health literacy, stigma, and behavioral health service use: the case of Latinx and non-Latinx whites. J Racial Ethn Health Disparities 2019 Dec;6(6):1122-1130. [doi: 10.1007/s40615-019-00614-8] [Medline: 31327136]
13. Grieb SM, Platt R, Vazquez MG, Alvarez K, Polk S. Mental health stigma among Spanish-speaking Latinos in Baltimore, Maryland. J Immigr Minor Health 2023 Oct;25(5):999-1007. [doi: 10.1007/s10903-023-01488-z] [Medline: 37213041]
14. Reducing the economic burden of unmet mental health needs. The White House. 2022 May 31. URL: https://bidenwhitehouse.archives.gov/cea/written-materials/2022/05/31/reducing-the-economic-burden-of-unmet-mental-health-needs/ [accessed 2024-05-15]
15. Chavira DA, Golinelli D, Sherbourne C, et al. Treatment engagement and response to CBT among Latinos with anxiety disorders in primary care. J Consult Clin Psychol 2014 Jun;82(3):392-403. [doi: 10.1037/a0036365] [Medline: 24660674]
16. Escovar EL, Bocanegra ES, Craske MG, et al. Mediators of ethnic differences in dropout rates from a randomized controlled treatment trial among Latinx and non-Latinx White primary care patients with anxiety disorders. J Nerv Ment Dis 2023 Jun 1;211(6):427-439. [doi: 10.1097/NMD.0000000000001533] [Medline: 37252881]
17. Schmidt SA. Teletherapy used to breakdown access barriers. Telemed E Health 2024 May;30(5):1491-1494. [doi: 10.1089/tmj.2023.0612] [Medline: 38190285]
18. Cantor J, McBain RK, Kofner A, Hanson R, Stein BD, Yu H. Telehealth adoption by mental health and substance use disorder treatment facilities in the COVID-19 pandemic. Psychiatr Serv 2022 Apr 1;73(4):411-417. [doi: 10.1176/appi.ps.202100191] [Medline: 34407631]
19. McBain RK, Schuler MS, Breslau J, Kofner A, Wang L, Cantor JH. Telehealth availability for mental health care during and after the COVID-19 public health emergency. JAMA Netw Open 2024 Jul 1;7(7):e2420853. [doi: 10.1001/jamanetworkopen.2024.20853] [Medline: 38985472]
20. Higashi RT, Thakur B, Repasky EC, et al. Digital health technology use among Spanish speakers in the US: a scoping review. JAMA Netw Open 2025 May 1;8(5):e2510386. [doi: 10.1001/jamanetworkopen.2025.10386] [Medline: 40372754]

21. Martinez S, Sanchez Z. Connecting Spanish-speakers to mental health providers attuned to their culture. California Health Care Foundation. 2024 Feb 20. URL: https://www.chcf.org/resource/connecting-spanish-speakers-mental-health-providers-attuned-their-culture/ [accessed 2025-11-07]
22. Sanarai. 2025. URL: https://www.sanarai.com [accessed 2025-11-07]
23. Aguilar-Gaxiola S, Nuño M, Henderson S, et al. Licensed physicians from México pilot program: final report. : University of California, Davis; 2025 Mar 31 URL: https://health.ucdavis.edu/media-resources/crhd/documents/pdfs/lpmpp-final-report.pdf [accessed 2025-11-07]
24. Borghouts J, Eikey E, Mark G, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021 Mar 24;23(3):e24387. [doi: 10.2196/24387] [Medline: 33759801]
25. Lau CKY, Saad A, Camara B, Rahman D, Bolea-Alamanac B. Acceptability of digital mental health interventions for depression and anxiety: systematic review. J Med Internet Res 2024 Oct 28;26:e52609. [doi: 10.2196/52609] [Medline: 39466300]
26. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017 Jan 26;17(1):88. [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]
27. Perski O, Short CE. Acceptability of digital health interventions: embracing the complexity. Transl Behav Med 2021 Jul 29;11(7):1473-1480. [doi: 10.1093/tbm/ibab048] [Medline: 33963864]
28. Gulliver A, Calear AL, Sunderland M, Kay-Lambkin F, Farrer LM, Batterham PJ. Predictors of acceptability and engagement in a self-guided online program for depression and anxiety. Internet Interv 2021 Sep;25:100400. [doi: 10.1016/j.invent.2021.100400] [Medline: 34026569]
29. Deen TL, Fortney JC, Schroeder G. Patient acceptance of and initiation and engagement in telepsychotherapy in primary care. Psychiatr Serv 2013 Apr 1;64(4):380-384. [doi: 10.1176/appi.ps.201200198] [Medline: 23370530]
30. Catania J, Beaver S, Kamath RS, et al. Evaluation of digital mental health technologies in the United States: systematic literature review and framework synthesis. JMIR Ment Health 2024 Aug 30;11:e57401. [doi: 10.2196/57401] [Medline: 39213023]
31. Torous J, Linardon J, Goldberg SB, et al. The evolving field of digital mental health: current evidence and implementation issues for smartphone apps, generative artificial intelligence, and virtual reality. World Psychiatry 2025 Jun;24(2):156-174. [doi: 10.1002/wps.21299] [Medline: 40371757]
32. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs-principles and practices. Health Serv Res 2013 Dec;48(6 Pt 2):2134-2156. [doi: 10.1111/1475-6773.12117] [Medline: 24279835]
33. The R Project for Statistical Computing. 2021. URL: https://www.R-project.org [accessed 2025-11-07]
34. Dillon K, Duncan DS. Know you customers' "job to be done" is innovation inherently a hit-or-miss endeavor? not if you understand why customers make the choices they do. Harvard Business Review. 2016. URL: https://hbr.org/2016/09/know-your-customers-jobs-to-be-done [accessed 2020-11-05]
35. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
36. Braun V, Clarke V, Hayfield N, Davey L, Jenkinson E. Doing reflexive thematic analysis. In: Bager-Charleson S, McBeath A, editors. Supporting Research in Counselling and Psychotherapy: Palgrave Macmillan; 2022. [doi: 10.1007/978-3-031-13942-0_2]
37. Designated health professional shortages areas statistics. Health Resources and Services Administration. 2025. URL: https://data.hrsa.gov/default/generatehpsaquarterlyreport [accessed 2025-11-07]
38. Pro G, Brown C, Rojo M, Patel J, Flax C, Haynes T. Downward national trends in mental health treatment offered in Spanish: state differences by proportion of Hispanic residents. Psychiatr Serv 2022 Nov 1;73(11):1232-1238. [doi: 10.1176/appi.ps.202100614] [Medline: 35502519]
39. Santiago-Rivera AL, Altarriba J, Poll N, Gonzalez-Miller N, Cragun C. Therapists' views on working with bilingual Spanish–English speaking clients: a qualitative investigation. Prof Psychol: Res Pract. 2009;40(5):436-443. [doi: 10.1037/a0015933]
40. Villalobos BT, Bridges AJ, Anastasia EA, Ojeda CA, Hernandez Rodriguez J, Gomez D. Effects of language concordance and interpreter use on therapeutic alliance in Spanish-speaking integrated behavioral health care patients. Psychol Serv 2016 Feb;13(1):49-59. [doi: 10.1037/ser0000051] [Medline: 26349073]
41. Psychologists reaching their limits as patients present with worsening symptoms year after year: 2023 Practitioner Pulse Survey. American Psychological Association. 2023. URL: https://www.apa.org/pubs/reports/practitioner/2023-psychologist-reach-limits [accessed 2025-12-04]
42. van Dijk DA, Meijer RM, van den Boogaard TM, Spijker J, Ruhé HG, Peeters FPML. Worse off by waiting for treatment? The impact of waiting time on clinical course and treatment outcome for depression in routine care. J Affect Disord 2023 Feb 1;322:205-211. [doi: 10.1016/j.jad.2022.11.011] [Medline: 36372129]
43. Dwight-Johnson M, Aisenberg E, Golinelli D, Hong S, O'Brien M, Ludman E. Telephone-based cognitive-behavioral therapy for Latino patients living in rural areas: a randomized pilot study. Psychiatr Serv 2011 Aug;62(8):936-942. [doi: 10.1176/ps.62.8.pss6208_0936] [Medline: 21807834]

44. Wendt D, Shafer K. Gender and attitudes about mental health help seeking: results from national data. Health Social Work 2016 Feb;41(1):e20-e28. [doi: 10.1093/hsw/hlv089]
45. Diep K, Frederiksen B, Long M, Ranji U, Salganicoff A. Access and coverage for mental health care: findings from the 2022 KFF women's health survey. KFF. 2022. URL: https://www.kff.org/mental-health/issue-brief/access-and-coverage-for-mental-health-care-findings-from-the-2022-kff-womens-health-survey [accessed 2025-12-04]
46. Aluisy A. Mental health–seeking behaviors for Hispanic individuals in the United States [Dissertation]. : University of South Florida; 2024 URL: https://digitalcommons.usf.edu/etd/10161 [accessed 2025-11-07]

Abbreviations

LPMPP: Licensed Physicians from Mexico Pilot Program
TA: thematic analysis

Edited by J Sarvestan; submitted 02.07.25; peer-reviewed by K Wells, S Mohanadas; revised version received 18.11.25; accepted 18.11.25; published 31.12.25.

Please cite as:
Arévalo Avalos MR, Fu J, Aguilera A, Suarez L
Reach, Engagement, and Acceptability of a Subclinical Telehealth Service for Spanish-Speaking Adults: Retrospective Mixed Methods Pilot Study
JMIR Form Res 2025;9:e80026
URL: https://formative.jmir.org/2025/1/e80026
doi:10.2196/80026

Original Paper

Impact of Cross-Sectoral Video Consultation on Perceived Care Coordination and Information Satisfaction in Cancer Care: Randomized Controlled Trial

Fereshteh Baygi[1], PhD; Theis Bitz Trabjerg[1], MD, PhD; Lars Henrik Jensen[2,3], MD, PhD; Maria Munch Storsveen[1], MSc; Sonja Wehberg[1], PhD; Jens Søndergaard[1], MD, PhD; Dorte Gilså Hansen[1], PhD

[1]Research Unit of General Practice, Department of Public Health, University of Southern Denmark, Odense, Denmark

[2]Lillebælt University Hospital, Department of Oncology, Vejle, Denmark

[3]Danish Colorectal Cancer Center South, Center of Clinical Excellence, Vejle Hospital, the Department of Regional Health Research, University of Southern Denmark, Vejle, Denmark

Corresponding Author:
Fereshteh Baygi, PhD
Research Unit of General Practice, Department of Public Health, University of Southern Denmark
Campusvej 55
Odense, 5230
Denmark
Phone: 45 65502348
Email: fbaygi@health.sdu.dk

Abstract

Background: Enhancing care coordination and sharing information in cancer care improves patient experiences by promoting clarity and satisfaction.

Objective: This study aims to assess the impact of cross-sectoral video consultation on patient perceptions of care coordination and satisfaction with received information compared to usual care.

Methods: This study presents secondary outcomes on patient perceptions of care coordination and satisfaction with received information from a 7-month follow-up of the Partnership Project. In this randomized controlled trial, patients with cancer were allocated to either an intervention group receiving cross-sectoral video consultation (oncologist, general practitioner, and patient) or a control group receiving usual care. Patients' perceptions of care coordination and information quality were assessed using the Australian Cancer Care Coordination Questionnaire (CCCQ) and the European Organisation for Research and Treatment of Cancer Quality of Life Information Questionnaire 25 at baseline and 7 months. Changes over time between groups were analyzed using generalized estimating equations.

Results: Of the 278 participants randomized (1:1), only 80 (28.8%) patients received the intervention due to technical and administrative issues. A total of 210 (75.5%) patients completed the baseline questionnaire, while 118 (42.4%) responded at 7 months. No significant differences were observed in the changes over time between the intervention and control groups in any outcome. The estimated differences in the change in score from baseline to 7 months were as follows: for the total CCCQ score, 1.11 (95% CI –2.32 to 4.53; *P*=.53); for the overall European Organisation for Research and Treatment of Cancer Quality of Life Information Questionnaire 25 score, 1.49 (95% CI –2.98 to 5.96; *P*=.51); for the CCCQ communication subscale, –1.49 (95% CI –1.33 to 4.31; *P*=.30); and for the navigation subscale, –0.03 (95% CI –1.52 to 1.46; *P*=.97).

Conclusions: Our findings indicate no statistically significant improvement in patients' reported care coordination or satisfaction with received information over 7 months. Technical issues with the video setup reduced fidelity rates and follow-up participation. Further research is needed to optimize the structure and content of cross-sectoral video consultations to better support patients' perceived outcomes.

Trial Registration: ClinicalTrials.gov NCT02716168; https://clinicaltrials.gov/study/NCT02716168

(JMIR Form Res 2025;9:e76910) doi:10.2196/76910

KEYWORDS

randomized controlled trials; video consultations; outcome assessment; care coordination; information satisfaction; cancer

Introduction

The cancer care system is complex, involving numerous transitions between primary and specialist care, often characterized by inadequate communication among health care providers [1]. Recent literature has noted that the frequent interactions and involvement of multiple care providers in cancer treatment contribute to its complexity, making it particularly suitable for telehealth-based coordination approaches [2]. This complexity underscores the necessity of coordination for patients who are currently receiving or who have received cancer treatment, ensuring that the health care system delivers high-quality cancer care and improves patient experiences [3]. Effective collaboration between different sectors is crucial for maintaining continuity of services, avoiding unnecessary overlap, and ensuring that care remains focused on the needs of patients [4].

Several studies, both randomized and nonrandomized, have been conducted to improve care coordination and continuity of care for patients with cancer [4-8]. Despite progress in cancer care research, the findings remain controversial [4,6]. For example, a systematic review found that most studies reported no significant changes for patients (eg, quality of life); providers, including general practitioners' (GPs) satisfaction with their role in patient care; or system outcomes (eg, frequency of GP visits). This review was unable to draw specific conclusions about the most effective models or interventions for improving cancer care coordination [6]. In contrast, a previous meta-analysis indicated that implementing cancer care coordination strategies led to improvements in most measured outcomes, such as overall patient experiences and quality of end-of-life care [4]. We believe this observed controversy highlights the challenges of developing effective interventions for the complex cancer care system, such as the mode of communication in care coordination and the degree of integration across sectors, which have not been addressed in previous studies [4].

None of the studies included in the previously mentioned systematic review [6] and meta-analyses [4] used a virtual intervention method to facilitate coordination. This gap is notable, particularly given the growing integration of digital health technologies into cancer care, including remote monitoring, patient portals, and electronic patient-reported outcomes, which have expanded significantly in recent years [9]. Among these, video consultation offers the potential to enable real-time, cross-sectoral dialogues [5]. Therefore, this study aimed to assess whether cross-sectoral video consultation can improve patient-perceived care coordination and information satisfaction over a 7-month period compared with usual care. We hypothesize that this approach will enhance these outcomes by facilitating timely, transparent, and inclusive communication across health care sectors, offering a potentially innovative solution to improve cancer care coordination.

Methods

Study Design

This study was a randomized controlled trial (RCT) titled the Partnership Project [10]. The protocol, details of this study, and findings on primary outcomes (eg, single-item global assessment of intersectoral cooperation) have been published previously [5,10,11]. This study presents secondary outcomes on patient perceptions of care coordination using the Australian Cancer Care Coordination Questionnaire (CCCQ) [12] and satisfaction with received information using the European Organisation for Research and Treatment of Cancer Quality of Life Information Questionnaire 25 (EORTC QLQ-INFO25) [13] from a 7-month follow-up survey in a shared video-based consultation.

Participants and Setting

Patients diagnosed with any type of cancer and starting their first chemotherapy treatment at the Department of Oncology, Lillebælt Hospital, University Hospital of Southern Denmark, were considered newly diagnosed and invited to participate in this study. The eligibility criteria included (1) being aged >18 years, (2) being proficient in speaking and reading Danish, and (3) having an oncologist's estimate of a survival time exceeding 7 months.

This trial was concluded upon reaching the predetermined sample size for patient inclusion.

Sample Size

The sample size was based on estimates from a previous Danish RCT aiming to improve GP involvement in cancer follow-up. To detect a clinically meaningful difference with 90% power and a 5% significance level, 194 participants were required. Considering an anticipated dropout rate of 30%, the recruitment goal was set at 278 patients. Detailed information on sample size has been published previously [5,11].

Intervention and Usual Care

Patients in the intervention group received a shared video consultation involving their GP, an oncologist, and themselves, in addition to usual care. These consultations were scheduled within 12 weeks of inclusion and could be held at either the GP's or oncologist's office based on patient preference. The oncologist typically chaired the session, supported by an oncology nurse, and both clinicians received a consultation guide beforehand.

The control group received usual care, which consisted of standard communication between the department of oncology and primary care, including electronic summary letters sent to the GP after each oncology visit, optional phone contact, and patient access to their GP.

A detailed description of the intervention and control groups, consultation structure, and implementation logistics has been published previously [5].

Randomization and Blinding

Patients were assigned in a 1:1 ratio through block randomization, with block sizes and sequences managed by

REDCap (Research Electronic Data Capture; Vanderbilt University) [14] data manager. A project nurse conducted the randomization and patient enrollment following informed consent. Allocation was transparent to patients, GPs, and oncologists, but blinding was maintained during baseline data collection. Data analysts remained blinded throughout. GPs of control group patients were not formally informed until they received the survey. Detailed information can be found in our previous publication [5].

Outcomes and Instruments

Patients were asked to complete questionnaires at baseline and at 4 and 7 months. Upon arrival at the department of oncology, they received information about the study, their perceived role, and the possibility of leaving the study as well as a consent form and a paper-based baseline questionnaire, which was collected by an outpatient nurse after enrollment. Follow-up questionnaires were distributed electronically via REDCap [14]. An overview of the secondary outcomes in this study is provided in Multimedia Appendix 1. The outcomes included items and subscales from CCCQ [12] (22 items) and EORTC QLQ-INFO25 [13] (25 items).

For this study, the CCCQ was translated from English to Danish following the guidelines for cross-cultural adaptation of self-report measures proposed by Beaton et al [15]. The translation procedures were documented in separate reports using the template developed by the American Academy of Orthopaedic Surgeons and recommended by Beaton et al [15].

The CCCQ evaluates patients' perceptions of cancer care coordination over a recall period of 3 months [12]. The CCCQ consists of 22 items: 2 global items, including global rating of coordination of care (item 21) and global rating of the quality of the received care (item 22) and 20 items that form 2 subscales, including the communication subscale (1-13; ranging from 13 to 65) and navigation subscale (items 14-20; ranging from 7 to 35). Responses are provided on 5-point Likert scales, ranging from "never" (1) to "always" (5), except for the global items, which use a 10-point Likert scale, ranging from "very poor" (1) to "very good" (10).

EORTC QLQ-INFO25, consisting of 25 items, assesses various aspects of patient satisfaction with the information provided during cancer treatment. Although no specific recall period was defined, patients were asked to consider their current cancer trajectory. The questionnaire includes a total score, 4 subscales covering information about the disease (items 1-4), medical tests (items 5-7), treatment (items 8-13) and other services (items 14-17) as well as 8 single items (items 18-25) addressing information about different care locations, self-help resources, written material, audiovisual formats, satisfaction with the information received, the need for more or less information, and the overall helpfulness of the information. In total, 4 additional yes-or-no questions explored the formats and quantity of information received by participants. Specifically, they asked whether participants had received written materials or audiovisual content (eg, CD, tape, or video), whether they wished to receive information (with an open-ended follow-up on preferred topics), and whether they felt they had received too much information (also with an option to specify topics). Responses to the core items were given on a 4-point Likert scale, ranging from "not at all" (1) to "very much" (4). All scores were linearly transformed to a scale ranging from 0 to 100 according to the scoring manual [13]. For the 4 items (written information, item 20; information on CD tape/video, item 21; wish to receive more information, item 23; and wish to receive less information, item 24) that have dichotomous response options (yes=1 or no=0), scores were treated as binary indicators.

Subscale score calculations were performed as follows:

- For CCCQ, subscale scores were also calculated as the mean of the items within each subscale. For Communication, up to 7 items could be missing, and for Navigation, up to 3 items could be missing. Scores of missing items were imputed using the mean of the nonmissing items within the respective subscale.
- For the EORTC QLQ-INFO25, subscale scores were calculated as the mean of the items within each subscale. For information about the disease, up to 2 single-item scores could be missing for the subscale score to remain valid. For information about medical tests, up to 1 item score could be missing; for information about treatments, up to 3 item scores could be missing; for information about other services, up to 2 item scores could be missing; for information about things you can do to help yourself, up to 13 scores could be missing. The remaining items were single items.

The direction of the answer scale varied by item. For the CCCQ Navigation subscale, a low score indicated a positive attitude toward the question, whereas for all other CCCQ subscales and single items, a high score indicated a positive attitude. For analysis purposes, all scores were aligned so that higher values indicated a positive attitude toward the questions. Specifically, item scores for items 14 to 20 were reversed coded to ensure consistent interpretation, such that an original score of 5 was recoded to 1, 4 to 2, 3 remained unchanged, 2 to 4, and 1 to 5.

Outcomes were measured at baseline and at 4 and 7 months after baseline. Coding was performed separately for each time point.

Other Parameters

Demographic data for patients, including age, gender, education, marital status, having children, employment status, comorbidities, and cancer diagnosis or type, were collected via a questionnaire completed by patients at baseline.

Statistical Analysis

A deviation from the initial analysis strategy was due to substantial missing data across all outcome variables at the 7-month follow-up (refer to the sample size reported for each outcome variable in the Results section). Detailed information on the original statistical analyses can be found in the published protocol [10] and our previous publication [5]. The revised statistical analysis is described subsequently.

For each secondary outcome, we compared the change from baseline to 7 months between the 2 groups using a linear regression model, following an intention-to-treat principle. The

model, applied to measurements at both baseline and 7 months, used generalized estimating equations to account for within-patient clusters, with robust variance estimation. The group difference was modeled as a time-by-group interaction. No additional covariates were included in this analysis. We assumed that missing data were missing completely at random, meaning that the probability of missingness was unrelated to unmeasured factors after accounting for treatment group and baseline outcome. The generalized estimating equation approach ensured robustness in the presence of missing data. In addition, we calculated the individual changes from baseline to follow-up, restricted to complete cases, and presented Cohen *d* as an effect size measure and *P* values based on a 2-sample 2-tailed *t* test (Multimedia Appendix 2).

Data analyses were performed using Stata (version 18; StataCorp) [16], with a significance level set at 5% ($P<.05$).

Ethical Considerations

Statement Regarding Human Participant Research Ethics Review

Ethics approval was obtained from the Regional Ethics Committee on Biomedical Research in Denmark (S-20142000-138) and the Danish Data Protection Agency (2014-41-3534).

Informed Consent

Informed consent was obtained from patients at the Department of Oncology, Vejle Hospital, Denmark. Consent covered participation in the RCT, video recordings, and patient-reported assessments. Consent forms were securely stored at the Clinical Research Unit, Vejle Hospital. As the unit of randomization was the patient, consent from GPs was not legally required. However, oral consent was obtained from GPs whose patients were allocated to the intervention group, and written information about the study was distributed to all GPs in the Region of Southern Denmark before recruitment. GPs who declined participation were excluded from this study.

Privacy and Confidentiality Protection

All video consultations were conducted using the Region of Southern Denmark's secure videoconferencing infrastructure, which ensured encrypted communication and excluded any involvement of third-party data processing. The intervention guide provided to oncologists and GPs included instructions for managing potentially sensitive patient disclosure with care and professionalism.

Compensation

Patients and oncologists did not receive financial compensation. GPs were reimbursed through the region's standard payment system for participating in video consultations and completing questionnaires, in line with cross-sector cooperation agreements.

Further details on the ethical framework and implementation are available in our previous publication [5].

Results

Recruitment and Baseline Data

Patients were enrolled in this study between June 2016 and November 2019. A total of 278 patients were randomly assigned to the intervention (n=139, 50%) and control (n=139, 50%) groups. However, due to GP-related issues (n=22, 15.8%), administrative (n=8, 5.8%) and technical issues (n=15, 10.8%), and clinical (n=3, 2.2%) or patient-related issues (n=8, 5.8%), only 80 (57.6%) patients received the intervention as intended. Figure 1 shows the participation flowchart.

Figure 1. Flowchart of participant enrollment and allocation in a randomized controlled trial on cross-sectoral video consultation in cancer care in Denmark (between June 2016 and November 2019).

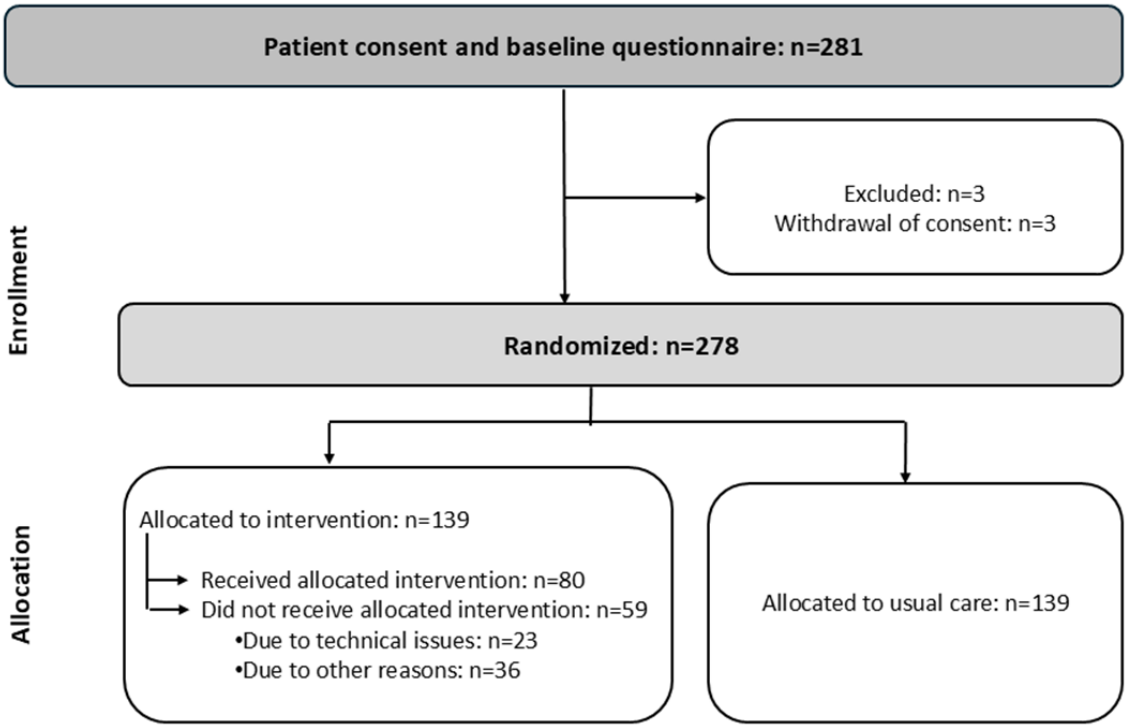


Table 1 presents the baseline characteristics of patients in the intervention and control groups. Patients in both groups exhibited similar baseline characteristics; however, comorbidities were more prevalent in the control group (n=81, 58.3%) compared to the intervention group (n=65, 46.8%).

Table 1. Baseline characteristics of patients in the intervention and control groups.

Characteristic	All (n=278)	Control group (n=139)	Intervention group (n=139)
Age (y), mean (SD)	65.2 (10.6)	63.8 (11.0)	66.6 (10.0)
Sex, n (%)			
Male	155 (55.8)	77 (55.4)	78 (56.1)
Female	123 (44.2)	62 (44.6)	61 (43.9)
Education, n (%)			
Primary and upper secondary school	176 (63.3)	85 (61.2)	88 (63.3)
Further education (3-4 y)	76 (27.3)	41 (29.5)	35 (25.2)
Higher education (≥5 y)	16 (5.8)	7 (5.0)	9 (6.5)
Marital status, n (%)			
Single or information missing[a]	81 (29.1)	48 (34.5)	33 (23.7)
Married or residing with a companion	197 (70.9)	91 (65.5)	106 (76.3)
Children living at home, n (%)			
No children at home or information missing[a]	244 (87.8)	120 (86.3)	124 (89.2)
Children at home	34 (12.2)	19 (13.7)	15 (10.8)
Work status, n (%)			
Employed	89 (32)	46 (33.1)	43 (30.9)
Public benefits	15 (5.4)	9 (6.5)	6 (4.3)
Retired or information missing[a]	174 (62.6)	84 (60.4)	90 (64.7)
Comorbidity, n (%)			
No	132 (47.5)	58 (41.7)	74 (53.2)
Yes	146 (52.5)	81 (58.3)	65 (46.8)
Diagnosis or cancer type, n (%)			
Breast	33 (11.9)	17 (12.2)	16 (11.5)
Gynecologic	13 (4.7)	4 (2.9)	9 (6.5)
Lung	106 (38.1)	53 (38.1)	53 (38.1)
Gastrointestinal	110 (39.6)	56 (40.3)	54 (38.8)
Other	16 (5.8)	9 (6.5)	7 (5)
Incident cancer (yes or information missing[a])	255 (91.7)	126 (90.6)	129 (92.8)

[a]There were fewer than 3 patients with missing information on marital status, the number of children at home, or work status and 6 patients with missing information on cancer incidents. These patients were grouped with the indicated categories.

Multimedia Appendix 3 provides an overview of the missing data regarding patient care cooperation (CCCQ subscales) and information outcomes (EORTC QLQ-INFO25) at various time points for both the control and intervention groups. While the number and proportion of missing data vary across each variable, a consistent pattern was observed across all variables, where missing data were minimal at baseline but increased substantially at the 4- and 7-month follow-ups.

Outcomes and Estimations

Table 2 presents patients' perceptions of care coordination and satisfaction with the information received between the primary sector and the department of oncology. No statistically significant differences were observed in the changes over time between the intervention and control groups in any outcome.

Table 2. Patients' perceptions of care coordination and satisfaction with the information received across the primary sector and the department of oncology in a randomized controlled trial in Denmark (between June 2016 and November 2019).[a]

Outcomes and group	Baseline		7 mo		Estimated change (95% CI)	Group-time interaction (95% CI)	*P* value
	n	Mean (SD)	n	Mean (SD)			
Australian Cancer Care Coordination Questionnaire : subscores and single items							
Global rating of coordination of care							
Control	134	8.37 (1.55)	59	8.59 (1.55)	0.14 (–0.23 to 0.51)	—[b]	—
Intervention	136	8.52 (1.40)	67	8.73 (1.67)	0.13 (–0.29 to 0.54)	–0.01 (–0.57 to 0.54)	.96
Global rating of the quality of received care							
Control	136	8.74 (1.24)	59	9.08 (1.19)	0.21 (–0.06 to 0.48)	—	—
Intervention	135	8.86 (1.15)	67	8.91 (1.54)	–0.02 (–0.38 to 0.35)	–0.22 (–0.68 to 0.23)	.33
Communications							
Control	135	46.79 (8.48)	59	47.80 (8.40)	0.16 (–2.11 to 1.80)	—	—
Intervention	134	48.13 (8.97)	68	50.06 (8.46)	–1.34 (0.70 to 3.37)	–1.49 (–1.33 to 4.31)	.30
Navigation							
Control	135	30.30 (3.80)	59	30.36 (4.78)	–0.43 (–1.49 to 0.62)	—	—
Intervention	135	30.79 (3.83)	67	30.63 (4.94)	–0.47 (–1.52 to 0.59)	–0.03 (–1.52 to 1.46)	.97
Total							
Control	135	77.08 (10.79)	59	78.16 (11.33)	–0.84 (–3.27 to 1.59)	—	—
Intervention	133	60.42 (8.08)	67	58.60 (8.50)	0.26 (–2.15 to 2.67)	1.11 (–2.32 to 4.53)	.53
European Organisation for Research and Treatment Quality of Life Information Questionnaire 25 : subscores and single items							
Information about the disease							
Control	135	59.75 (22.04)	59	65.77 (22.37)	4.09 (–1.82 to 10.00)	—	—
Intervention	134	58.89 (20.83)	68	67.03 (24.75)	5.96 (0.47 to 11.44)	1.87 (–6.19 to 9.93)	.65
Information about medical tests							
Control	135	76.09 (22.31)	60	77.22 (23.89)	–1.38 (−7.09 to 4.32)	—	—
Intervention	133	81.29 (18.62)	68	80.23 (23.34)	–2.23 (–7.49 to 3.02)	–0.85 (–8.61 to 6.91)	.83
Information about treatment							
Control	134	65.17 (21.44)	60	65.54 (21.20)	–2.62 (–7.77 to 2.52)	—	—
Intervention	133	67.54 (20.22)	68	69.22 (24.45)	0.39 (–5.34 to 6.13)	3.02 (–4.69 to 10.72)	.44
Information about other services							
Control	132	40.70 (26.31)	60	46.67 (28.01)	2.96 (–2.90 to 8.82)	—	—
Intervention	133	42.88 (26.23)	68	54.62 (30.60)	9.24 (2.43 to 16.04)	6.27 (–2.71 to 15.25)	.17
Information about different places of care							
Control	129	37.21 (36.24)	59	44.07 (33.01)	1.75 (–6.40 to 9.90)	—	—
Intervention	131	36.90 (36.81)	67	47.76 (35.87)	10.49 (1.09 to 19.90)	8.74 (–3.70 to 21.19)	.17

Outcomes and group	Baseline		7 mo		Estimated change (95% CI)	Group-time interaction (95% CI)	*P* value
	n	Mean (SD)	n	Mean (SD)			
Information about things you can do to help yourself							
Control	131	55.47 (32.45)	59	58.76 (31.16)	–0.49 (–7.04 to 6.07)	—	—
Intervention	128	64.58 (29.80)	66	59.60 (33.34)	–7.50 (–15.23 to 0.22)	–7.02 (–17.15 to 3.11)	.17
Satisfaction with the information received							
Control	136	83.33 (23.31)	59	80.79 (24.14)	–4.25 (–10.25 to 1.75)	—	—
Intervention	136	90.20 (17.27)	68	82.84 (22.67)	–6.81 (–11.65 to –1.97)	–2.56 (–10.27 to 5.15)	.52
Overall, the information has been helpful							
Control	134	84.33 (21.51)	60	78.33 (24.41)	–7.57 (–13.20 to –1.93)	—	—
Intervention	135	89.38 (15.59)	68	84.31 (19.51)	–5.38 (–9.85 to –0.90)	2.19 (–5.01 to 9.38)	.55
Written information							
Control	135	91.11 (28.56)	60	88.33 (32.37)	–3.27 (–11.61 to 5.08)	—	—
Intervention	134	88.81 (31.65)	67	80.60 (39.84)	–8.69 (–19.53 to 2.15)	–5.42 (–19.10 to 8.26)	.44
Information on CD, tape, or video							
Control	134	3.73 (19.02)	57	3.51 (18.56)	0.71 (–3.19 to 4.61)	—	—
Intervention	135	4.44 (20.68)	65	1.54 (12.40)	–2.29 (–5.11 to 0.52)	–3.01 (–7.82 to 1.80)	.22
Wish to receive more information							
Control	133	32.33 (46.95)	59	16.95 (37.84)	–15.08 (–26.39 to –3.76)	—	—
Intervention	136	24.26 (43.03)	64	12.50 (33.33)	–12.13 (–22.46 to –1.79)	2.95 (–12.37 to 18.27)	.71
Global score							
Control	122	52.30 (12.96)	53	51.13 (13.81)	–3.53 (–6.34 to –0.72)	—	—
Intervention	124	54.62 (10.63)	55	52.94 (14.90)	–2.04 (–5.52 to 1.44)	1.49 (–2.98 to 5.96)	.51

[a]The efficacy of the intervention was evaluated in unadjusted generalized estimating equation population-averaged linear regression models following an intention-to-treat approach, where the effectiveness was estimated as a group-time interaction term. For the control group, this within-group change was the estimated regression coefficient for time (follow-up vs baseline), while for the intervention group, the within-group change was based on the estimated coefficient for time and the estimated group-time interaction. The group effect was modeled as the difference between the within-group changes over time, that is, the coefficient of the group-time interaction term.

[b]Not applicable

The estimated within-group changes between baseline and 7-month follow-up for the total CCCQ score were –0.84 (95% CI –3.27 to 1.59; *P*=.498) in the control group and 0.26 (95% CI –2.15 to 2.67; *P*=.83) in the intervention group. The between-group difference was estimated as 1.11 (95% CI –2.32 to 4.53; *P*=.53). The estimated within-group changes between baseline and 7-month follow-up for the overall INFO25 score were –3.53 (95% CI –6.34 to –0.72; *P*=.01) in the control group and –2.04 (95% CI –5.52 to 1.44; *P*=.25) in the intervention group. The between-group difference was estimated as 1.49 (95% CI –2.98 to 5.96; *P*=.52). The estimated difference in the change from baseline to 7 months for the CCCQ communication subscale was –1.49 (95% CI –1.33 to 4.31; *P*=.30) and for the navigation subscale was –0.03 (95% CI –1.52 to 1.46; *P*=.97).

Discussion

Principal Findings

This study found that the addition of a cross-sectoral video consultation to usual care did not lead to statistically significant

improvements in patients' perceptions of care coordination and satisfaction with the information received over 7 months. There were no significant differences in changes over time between the control and intervention groups for any of the secondary outcomes.

Comparison to Prior Work

Telemedicine, including video consultation and phone calls, has emerged as a crucial tool for maintaining the continuity of medical care for patients with various medical conditions during the COVID-19 pandemic [17]. However, our study on the innovative health care model was conducted before the COVID-19 pandemic and the implementation of standard video setups. Since then, the technical aspects of video-based communication in health care have significantly improved, making it challenging to compare our findings with post–COVID-19 pandemic studies that use these advanced protocols.

Despite this advancement, recent research highlights that the impact of telemedicine on patient outcomes remains nuanced [18,19]. For instance, a meta-analysis on the efficacy of telemedicine for outpatients found that video consultation was feasible but did not significantly outperform face-to-face care in oncology in terms of patient satisfaction and attendance [18]. Authors in an editorial with a conceptual perspective lens argued that while remote consultation may enhance access, they risk compromising rational continuity, information richness, and shared decision-making [19]. This underscores the importance of context-sensitive consultation approaches, guided by clinical judgment and patient preference rather than policy mandates [19]. We believe that while these 2 studies focus on different dimensions, both highlight limitations of telemedicine and reinforce the need for thoughtful implementation.

A systematic review that analyzed 51 studies across multiple medical disciplines, including primary care, oncology, dermatology, and so on, revealed that telemedicine provided valuable support to traditional medicine during the COVID-19 pandemic [17]. The review highlighted a high level of patient satisfaction with telemedicine, particularly in areas such as convenience, continuity of care, communication, and efficiency [17]. In contrast, our study, which focused on patients' perception of care coordination and satisfaction with the information received, found no significant improvement in these areas after the follow-up period. This discrepancy might be due to the different variables and methodologies used in each study as well as several limitations in our study, such as many missing values at follow-up, the lack of standard protocols for video consultation due to the timing of the trial, and other implementation challenges, including technical issues or scheduling difficulties, that were comprehensively explained in our previous publication [5].

The other study comparing the content and quality of various modes of consultation (eg, video, telephone, and face-to-face) found that face-to-face consultations enabled more information exchange between GPs and patients compared to video and telephone consultations [20]. The observed difference between our study and the aforementioned study highlights the potential limitations of video consultation in achieving patient outcomes and emphasizes the importance of face-to-face interaction in the health care setting. However, due to several limitations in our study that have been discussed in our previous publication [5], we believe that comparisons with other studies and the generalizability of our findings should be approached with caution. Specifically, the high rates of missing data in our study hinder us from drawing concrete conclusions about the intervention effect on patients' perception of care coordination and satisfaction with the information received. Instead, we encourage the research community to learn from our trial's challenges [5] and insights to improve future studies on this research area.

Limitations of This Study

First, a key limitation of this study concerns the use of the CCCQ questionnaire, which was translated from English into Danish for this project. Although the translation followed the established guidelines for cross-cultural adaptation of self-report measures [15], the Danish version had not been validated before this study. This may impact the reliability and cultural appropriateness of the instrument in the Danish context. Second, the considerable amount of missing data observed at the 7-month follow-up reduced statistical power, affecting the ability to detect significant effects. Third, another limitation is the likely ceiling effect observed in both the CCCQ and EORTC QLQ-INFO25. At baseline, scores in both groups were high, leaving limited room for measurable improvement, potentially masking the intervention's impact. Implementation challenges further limited this study. These included low completion rates of video consultations, technical and administrative issues, and the absence of a standardized video setup during the trial—details previously reported [5]. In addition, the absence of structured fidelity metrics, such as session counts, duration, and quality, limits the interpretability of intervention delivery. Although descriptive data on session uptake were collected, logistical challenges and a large amount of missing data prevented meaningful statistical analyses. Finally, the absence of statistically significant findings likely reflects a combination of factors, including limited intervention uptake, variability in participant engagement, measurement limitations (eg, the use of translated but unvalidated questionnaires), and inconsistent implementation. These findings highlight the need for future studies to use validated outcome measures and incorporate robust implementation and fidelity-tracking strategies.

Implications and Future Directions

From the clinical and implementation perspective, the findings suggest that when baseline care is already perceived as high quality, additional interventions may offer limited incremental value. Future efforts to improve care coordination and patient information may be more appropriately targeted toward settings or populations where baseline performance is suboptimal, as there may be more potential for meaningful improvement in these contexts. In addition, future studies should incorporate structured fidelity metrics (eg, session completion rate) to better understand intervention delivery and identify implementation barriers. This would support more adaptation to clinical settings and enhance scalability within established frameworks, such as the Consolidated Framework for Implementation Research and

the reach, effectiveness, adoption, implementation, and maintenance framework, which help guide and evaluate the adaptation, effectiveness, and sustainability of health interventions [21-23]. Furthermore, redesigning the intervention to address technical and administrative challenges, such as streamlining, scheduling, integrating automated reminders, and improving digital infrastructure, could enhance feasibility and uptake. Co-designing future iterations with end users may further improve usability and engagement, supporting more effective implementation and transferability to other settings.

Acknowledgments

Adherence to reporting standards was ensured using the CONSORT-HEALTH checklist (See Multimedia Appendix 4). During the preparation of this manuscript, the authors used Microsoft Copilot to assist with refinement, correction, editing, and formatting of the initial draft and subsequent revisions to improve clarity of language [24]. All artificial intelligence–assisted revisions were carefully reviewed and modified by the authors, who take full responsibility for the final version of the publication.

Funding

Financial support was provided by the Danish Cancer Society (R156-A10018-16-S49), the Region of Southern Denmark (14/39009), ML Jørgensen and Gunnar Hansen Foundation, and the University of Southern Denmark. The sponsors played no role in the planning, design, or implementation of the study.

Data Availability

Data collection was conducted in compliance with Article 6 the General Data Protection Regulation, based on participants' consent. This consent does not authorize public disclosure of the data; therefore, the dataset cannot be shared with unauthorized parties.

Authors' Contributions

TBT and LHJ conceived the study, participated in study design and data collection, and revised the manuscript. JS and DGH conceived the study, participated in study design, and revised the manuscript critically. MMS and SW analyzed the data, participated in the interpretation of the results, and revised the manuscript. FB made substantial contributions to the interpretation of the results and wrote the manuscript. All authors reviewed and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Overview of secondary outcomes.
[DOCX File , 32 KB - formative_v9i1e76910_app1.docx]

Multimedia Appendix 2
Individual differences for complete cases.
[DOCX File , 21 KB - formative_v9i1e76910_app2.docx]

Multimedia Appendix 3
Overview of the missing data for patient's care cooperation (CCCQ) and satisfaction of information received (EORTC QLQ-INFO25) as secondary outcomes in both control and intervention groups at baseline and follow-up.
[DOCX File , 40 KB - formative_v9i1e76910_app3.docx]

Multimedia Appendix 4
CONSORT checklist.
[PDF File (Adobe PDF File), 162 KB - formative_v9i1e76910_app4.pdf]

References

1. Clauser SB, Wagner EH, Aiello Bowles EJ, Tuzzio L, Greene SM. Improving modern cancer care through information technology. Am J Prev Med 2011 May;40(5 Suppl 2):S198-S207 [FREE Full text] [doi: 10.1016/j.amepre.2011.01.014] [Medline: 21521595]

2. Rendle KA, Tan AS, Spring B, Bange EM, Lipitz-Snyderman A, Morris MJ, et al. A framework for integrating telehealth equitably across the cancer care continuum. J Natl Cancer Inst Monogr 2024 Jun 26;2024(64):92-99. [doi: 10.1093/jncimonographs/lgae021] [Medline: 38924790]
3. Weaver SJ, Jacobsen PB. Cancer care coordination: opportunities for healthcare delivery research. Transl Behav Med 2018 May 23;8(3):503-508 [FREE Full text] [doi: 10.1093/tbm/ibx079] [Medline: 29800404]
4. Gorin SS, Haggstrom D, Han PK, Fairfield KM, Krebs P, Clauser SB. Cancer care coordination: a systematic review and meta-analysis of over 30 years of empirical studies. Ann Behav Med 2017 Aug;51(4):532-546. [doi: 10.1007/s12160-017-9876-2] [Medline: 28685390]
5. Baygi F, Bitz Trabjerg T, Jensen LH, Munch Storsveen M, Wehberg S, Sisler JJ, et al. Challenges of cross-sectoral video consultation in cancer care on patients' perceived coordination: randomized controlled trial. JMIR Cancer 2025 Feb 11;11:e60158 [FREE Full text] [doi: 10.2196/60158] [Medline: 39933723]
6. Tomasone JR, Brouwers MC, Vukmirovic M, Grunfeld E, O'Brien MA, Urquhart R, et al. Interventions to improve care coordination between primary healthcare and oncology care providers: a systematic review. ESMO Open 2016;1(5):e000077 [FREE Full text] [doi: 10.1136/esmoopen-2016-000077] [Medline: 27843639]
7. Wulff CN, Vedsted P, Søndergaard J. A randomized controlled trial of hospital-based case management in cancer care: a general practitioner perspective. Fam Pract 2013 Feb;30(1):5-13. [doi: 10.1093/fampra/cms050] [Medline: 22952209]
8. Sprague BL, Dittus KL, Pace CM, Dulko D, Pollack LA, Hawkins NA, et al. Patient satisfaction with breast and colorectal cancer survivorship care plans. Clin J Oncol Nurs 2013 Jun;17(3):266-272 [FREE Full text] [doi: 10.1188/13.CJON.17-03AP] [Medline: 23722604]
9. Haemmerle R, Paludo J, Haddad TC, Pritchett JC. The growing role of digital health tools in the care of patients with cancer: current use, future opportunities, and barriers to effective implementation. Curr Oncol Rep 2024 Jun;26(6):593-600. [doi: 10.1007/s11912-024-01534-5] [Medline: 38652424]
10. Trabjerg TB, Jensen LH, Søndergaard J, Sisler JJ, Hansen DG. Improving continuity by bringing the cancer patient, general practitioner and oncologist together in a shared video-based consultation - protocol for a randomised controlled trial. BMC Fam Pract 2019 Jun 25;20(1):86 [FREE Full text] [doi: 10.1186/s12875-019-0978-8] [Medline: 31238886]
11. Trabjerg TB, Jensen LH, Søndergaard J, Sisler JJ, Hansen DG. Cross-sectoral video consultations in cancer care: perspectives of cancer patients, oncologists and general practitioners. Support Care Cancer 2021 Jan;29(1):107-116 [FREE Full text] [doi: 10.1007/s00520-020-05467-0] [Medline: 32318872]
12. Young JM, Walsh J, Butow PN, Solomon MJ, Shaw J. Measuring cancer care coordination: development and validation of a questionnaire for patients. BMC Cancer 2011 Jul 15;11:298 [FREE Full text] [doi: 10.1186/1471-2407-11-298] [Medline: 21756360]
13. Arraras JI, Greimel E, Sezer O, Chie W, Bergenmar M, Costantini A, et al. An international validation study of the EORTC QLQ-INFO25 questionnaire: an instrument to assess the information given to cancer patients. Eur J Cancer 2010 Oct;46(15):2726-2738. [doi: 10.1016/j.ejca.2010.06.118] [Medline: 20674333]
14. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
15. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine (Phila Pa 1976) 2000 Dec 15;25(24):3186-3191. [doi: 10.1097/00007632-200012150-00014] [Medline: 11124735]
16. What's new in STATA 18. STATA. URL: https://www.stata.com/stata18/ [accessed 2025-02-03]
17. Pogorzelska K, Chlabicz S. Patient satisfaction with telemedicine during the COVID-19 pandemic-a systematic review. Int J Environ Res Public Health 2022 May 17;19(10):6113 [FREE Full text] [doi: 10.3390/ijerph19106113] [Medline: 35627650]
18. Uemoto Y, Yamanaka T, Kataoka Y, Wada Y, Aoyama Y, Kizawa R, et al. Efficacy of telemedicine using videoconferencing systems in outpatient care for patients with cancer: a systematic review and meta-analysis. JCO Clin Cancer Inform 2022 Nov;6:e2200084. [doi: 10.1200/CCI.22.00084] [Medline: 36417685]
19. Greenhalgh T, Rosen R. Remote by default general practice: must we, should we, dare we? Br J Gen Pract 2021 Apr;71(705):149-150 [FREE Full text] [doi: 10.3399/bjgp21X715313] [Medline: 33771790]
20. Hammersley V, Donaghy E, Parker R, McNeilly H, Atherton H, Bikker A, et al. Comparing the content and quality of video, telephone, and face-to-face consultations: a non-randomised, quasi-experimental, exploratory study in UK primary care. Br J Gen Pract 2019 Sep;69(686):e595-e604 [FREE Full text] [doi: 10.3399/bjgp19X704573] [Medline: 31262846]
21. Reardon CM, Damschroder LJ, Ashcraft LE, Kerins C, Bachrach RL, Nevedal AL, et al. The Consolidated Framework for Implementation Research (CFIR) user guide: a five-step guide for conducting implementation research using the framework. Implement Sci 2025 Aug 16;20(1):39 [FREE Full text] [doi: 10.1186/s13012-025-01450-7] [Medline: 40819067]
22. Glasgow RE, Harden SM, Gaglio B, Rabin B, Smith ML, Porter GC, et al. RE-AIM planning and evaluation framework: adapting to new science and practice with a 20-year review. Front Public Health 2019;7:64 [FREE Full text] [doi: 10.3389/fpubh.2019.00064] [Medline: 30984733]

23. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health 1999 Sep;89(9):1322-1327. [doi: 10.2105/ajph.89.9.1322] [Medline: 10474547]
24. Recommendations for a classification of AI use in academic manuscript preparation. The International Association of Scientific, Technical & Medical Publishers. 2025. URL: https://stm-assoc.org/document/recommendations-for-a-classification-of-ai-use-in-academic-manuscript-preparation/ [accessed 2025-05-29]

Abbreviations

CCCQ: Cancer Care Coordination Questionnaire
EORTC QLQ-INFO25: European Organisation for Research and Treatment Quality of Life Information Questionnaire 25
GP: general practitioner
RCT: randomized controlled trial
REDCap: Research Electronic Data Capture

Edited by A Mavragani, J Sarvestan; submitted 04.05.25; peer-reviewed by S Bestavros; comments to author 01.08.25; revised version received 10.11.25; accepted 11.11.25; published 31.12.25.

Please cite as:
Baygi F, Trabjerg TB, Jensen LH, Munch Storsveen M, Wehberg S, Søndergaard J, Gilså Hansen D
Impact of Cross-Sectoral Video Consultation on Perceived Care Coordination and Information Satisfaction in Cancer Care: Randomized Controlled Trial
JMIR Form Res 2025;9:e76910
URL: https://formative.jmir.org/2025/1/e76910
doi:10.2196/76910
PMID:

Original Paper

Ghana Heart Initiative Training for Cardiac Arrest Management Among Health Care Professionals: Outcomes Evaluation Study (2019-2024)

Alfred Doku[1,2*], MPH, MBChB; Lawrence Sena Tuglo[3*], MPH; Chiedozie Osuoji[4*], MBBS, MSci; Juliette Edzeame[5], BPharm, MPH; Marisa Broni[5], BPharm; David Danso Mainoo[5*], MSci; Alberta Ewuziwaa Acquah[5*], MSci; Kwatetso Honny[5*], BPharm, MPH; Ron J G Peters[6*], MD, PhD; Charles Agyemang[2,7], PhD

[1]Department of Medicine and Therapeutics, University of Ghana Medical School, University of Ghana, Accra, Ghana

[2]Department of Public & Occupational Health, University of Amsterdam Medical Centre, University of Amsterdam, Amsterdam, Netherlands, The Netherlands

[3]Department of Nutrition and Dietetics, School of Allied Health Sciences, University of Health and Allied Sciences, Ho, Ghana

[4]Code Red Emergency Medical Services, Accra, Ghana

[5]Aya Integrated Healthcare Initiative, International Services, Deutsche Gesellschaft fur Internationale Zusammenarbeit (GIZ), Accra, Ghana

[6]Department of Cardiology, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands

[7]Department of Medicine, Johns Hopkins University School of Medicine, Johns Hopkins University, Baltimore, MD, United States

[*]these authors contributed equally

Corresponding Author:
Alfred Doku, MPH, MBChB
Department of Medicine and Therapeutics
University of Ghana Medical School
University of Ghana
Korle Bu Campus
Legon
Accra, LG 25
Ghana
Phone: 1 0244273573
Email: dokukavin@gmail.com

Abstract

Background: Health care professionals must stay updated with the latest guidelines for basic life support (BLS) and advanced cardiac life support (ACLS) to effectively assist patients during cardiac emergencies. Since its launch in 2018, the Ghana Heart Initiative has significantly enhanced the skills and knowledge of health care professionals in managing cardiovascular diseases, including cardiac emergencies.

Objective: This study aims to assess the knowledge and skills of BLS and ACLS among health care professionals immediately after training in Ghana.

Methods: This cross-sectional, training-based study involved 541 and 302 health care professionals trained in BLS and ACLS, respectively. Among them, 229 BLS and 124 ACLS-trained participants completed the questionnaires immediately after the training, and their data were included in the final analysis. Knowledge was assessed using a standardized questionnaire and an instructor-led skills evaluation based on the updated 2018 and 2020 American Heart Association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care.

Results: This study shows that 74.6% (171/229) of the health care professionals had adequate knowledge and skills in BLS. Those working in tertiary health care facilities were 80% less likely (adjusted odds ratio [AOR] 0.20, 95% CI 0.07-0.59; *P*=.003) to have adequate BLS knowledge and skills than those in primary health care facilities. Health care professionals from regions such as Volta and Oti were 4.94 times more likely to have adequate BLS knowledge and skills compared to those from Bono East (AOR 4.94, 95% CI 1.17-20.80; *P*=.03). Over 73.3% (91/124) of health care professionals had adequate knowledge and skills in ACLS. Males were 7.05 times more likely (AOR 7.05, 95% CI 2.69-18.46; *P*<.001) than females to possess adequate ACLS knowledge and skills.

Conclusions: Given an opportunity to learn and practice, health care professionals in Ghana attain adequate knowledge and skills in BLS and ACLS.

(JMIR Form Res 2025;9:e75536) doi:10.2196/75536

KEYWORDS

knowledge; skill assessment; basic life support; advanced cardiac life support; health professionals; Ghana

Introduction

Cardiac arrest, the sudden and complete shutdown of heart function, causes both respiratory and circulatory failure [1]. It is a critical emergency that demands immediate action [1]. Most adult cardiac arrests occur suddenly mainly due to heart diseases or systemic issues such as electrolyte and acid-base imbalances [2]. Therefore, delivering chest compression is essential for maintaining blood flow [3]. Prompt initiation of cardiopulmonary resuscitation (CPR) is crucial because it temporarily preserves life by ensuring that sufficient blood reaches vital organs such as the brain and heart [4]. Knowledge and skills in basic life support (BLS) and advanced cardiac life support (ACLS) are vital for managing sudden cardiac arrest, myocardial infarction, foreign body airway obstructions, respiratory failure [3], and other serious cardiovascular diseases (CVDs) [5,6].

The American Heart Association (AHA) plays a crucial role in emergency cardiovascular care (ECC) by regularly updating the CPR and ECC guidelines [7,8]. These guidelines provide a standard approach for delivering BLS and ACLS to revive patients [7,8]. BLS, which involves a series of care given to patients experiencing respiratory arrest, cardiac arrest, or airway obstruction, is a fundamental part of these guidelines [9]. The BLS course, based on AHA guidelines, teaches participants how to quickly recognize life-threatening emergencies, perform effective chest compressions, provide proper ventilation, and use an automated external defibrillator [9]. Similarly, the ACLS, also based on AHA guidelines, includes recognizing signs of sudden cardiac arrest, myocardial infarction, and complete airway obstruction, as well as performing CPR and defibrillation with an automated external defibrillator [1]. Following these guidelines, ACLS and CPR, when administered promptly, have been shown to improve survival rates in medical emergencies [1].

As health care professionals play a vital role in emergency response, they are expected to possess a strong understanding of the basic resuscitation principles in patient care and stay updated with the BLS and ACLS guidelines for rescuing patients who are unconscious or in cardiac arrest. However, studies conducted in low-income countries such as South Africa [10], Afghanistan [2], and Pakistan [11] have reported limited BLS knowledge and skills among health care professionals. Additionally, research in Ethiopia [1] and Spain [4] has shown insufficient understanding of ACLS among health care professionals (nurses and doctors) and nurses, respectively. Unlike in high-income countries, where strict licensing or certification is required before practice, few studies [1,2,9,11,12] have evaluated the BLS and ACLS knowledge of health care professionals in low-income countries.

In Ghana, there are no strict licensing requirements for health care professionals to be trained in current BLS and ACLS guidelines or to obtain a certificate before practicing. However, through the Ghana Heart Initiative, some health care professionals in Ghana have received training in BLS and ACLS based on AHA-certified guidelines, filling this critical gap in the health system [13]. To our knowledge, no study has assessed the knowledge and skills in BLS and ACLS among health care professionals in Ghana. Therefore, this study aims to evaluate their knowledge and skills in BLS and ACLS through training, enabling them to manage cardiac arrest effectively in Ghana.

Methods

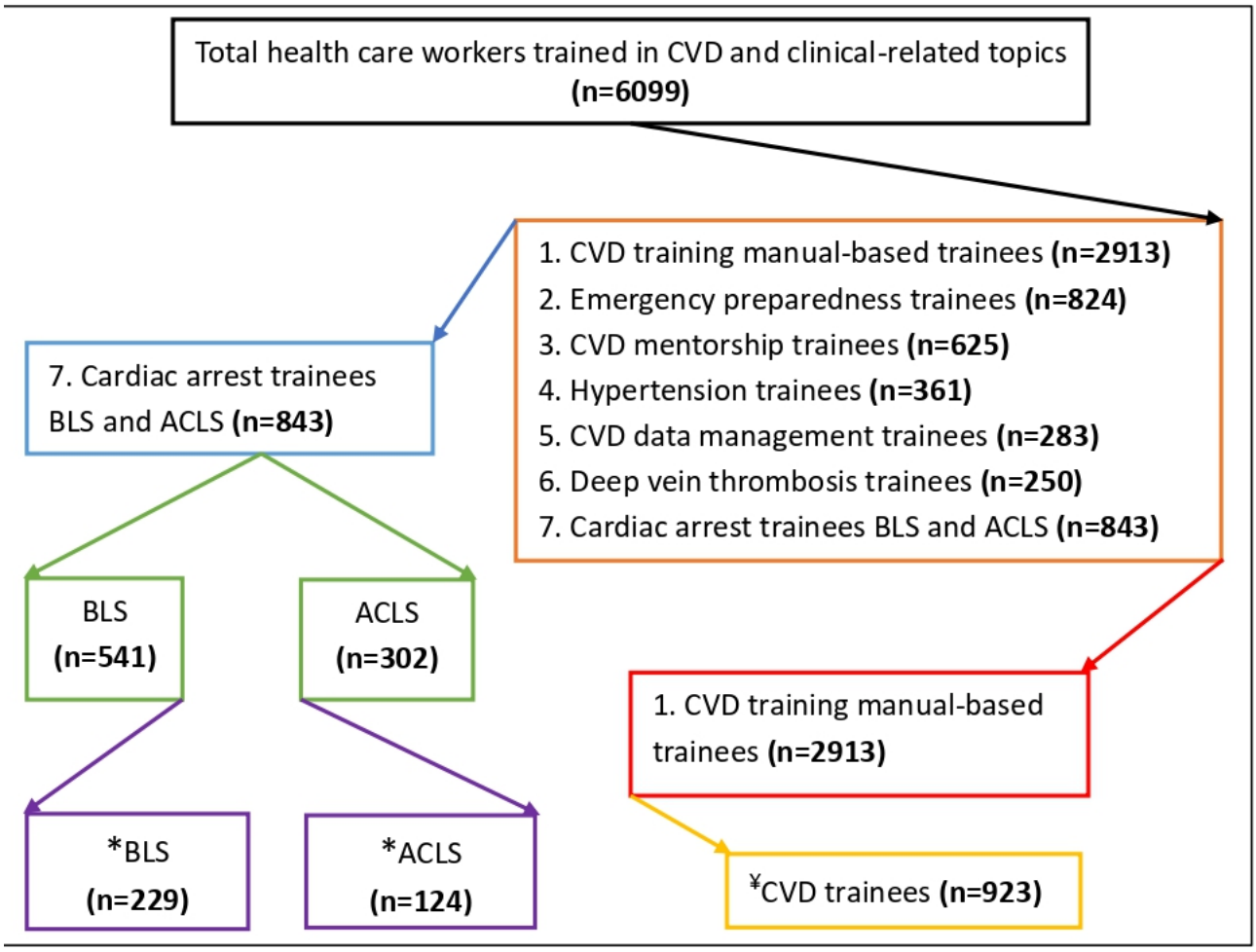
Study Setting and Recruitment

This study was conducted from 2019 to 2024 across 8 of Ghana's 16 administrative regions. These include Ashanti, Bono East, Central, Eastern, Greater Accra, Northern, Volta, and Oti. Ghana, located in West Africa, is classified as a lower-middle-income country. Ghana borders the Gulf of Guinea and the Atlantic Ocean, sharing borders with Côte d'Ivoire to the west, Burkina Faso to the north, and Togo to the east [14]. As of 2021, Ghana had an estimated population of 33 million, covering approximately 238,535 square kilometers. Accra serves as the capital and functions as the country's economic and administrative center [14]. To ensure fair representation, proportional sampling was used to select 44 health facilities, including primary, secondary, and tertiary levels, as well as both public and quasi-governmental facilities. These included 6 health centers, 27 secondary facilities, 5 tertiary hospitals, and 6 community health planning and services facilities [15]. Written memos were sent to the administrators of the selected health facilities, requesting at least 1 health professional such as a physician assistant, nurse, house officer, medical officer, senior medical officer, specialist, or trainee specialist to participate in the training. The sessions took place in the capital cities of these regions at designated locations to encourage high attendance among health care professionals.

Study Design and Population

This study employs a cross-sectional, training-based design involving 541 and 302 health care professionals recruited from selected health facilities and trained in BLS and ACLS, respectively. Among them, 229 BLS and 124 ACLS-trained participants completed the questionnaires immediately after the training, and their data were included in the final analysis (Figure 1).

Figure 1. Flow diagram of health care professionals trained in CVDs and clinical topics. *Completed the BLS and ACLS training questionnaires and included in the final analysis. ¥The details of CVD trainees and the data are presented in a different study. ACLS: advanced cardiac life support; BLS: basic life support; CVD: cardiovascular disease.



BLS and ACLS training

BLS and ACLS training took place over 1 and 3 days, respectively, conducted by Code Red, a certified training provider experienced in AHA. The course guide was based on the updated 2018 [7] and 2020 [8] AHA guidelines for CPR and ECC and was provided to participants at least 4 weeks prior to the training for preparation. The training packages included lectures with PowerPoint presentations, video materials, practical demonstrations, and hands-on exercises. The BLS course covered, among other topics, the chain of survival, the use of automated external defibrillators, team dynamics, CPR for infants and children, ventilation techniques, opioid-related emergencies, and choking. For ACLS, topics included, among others, cardiac arrest, acute coronary syndromes, stroke, arrhythmias, respiratory arrest, types of cardiac arrest (pulseless electrical activity, asystole, and ventricular tachycardia/fibrillation), and postcardiac arrest care.

Data Collection and Assessment

After the training, an assessment of knowledge was conducted immediately through the self-administration of a standardized questionnaire and CPR skills drills, according to the updated 2018 [7] and 2020 [8] AHA guidelines for CPR and ECC. The questionnaire included demographics (gender, profession, facility type, and region) as well as 20 multiple-choice questions on BLS and ACLS. Ten questions were on BLS, and the other 10 questions were on ACLS. The participants took at least 30-45 minutes immediately after the training to complete the questionnaire.

Classification of Knowledge and Skill Scores for BLS and ACLS

The knowledge and skill scores on BLS and ACLS were categorized into 2 levels. Scores ≥90% are considered adequate, while scores <90% are regarded as inadequate.

Statistical Analysis

Data were coded and entered into Microsoft Excel 2016. Data were imported into SPSS software (version 25; IBM Corp) for analysis. Categorical data were presented as frequencies and percentages. When appropriate, the chi-square test or Fisher exact test was used to determine the associations between the BLS and ACLS knowledge and skill scores and demographic data. A *P* value <.05 was considered significant, and only those variables were included in the logistic regression analyses to test for associations.

Ethical Considerations

Ethical clearance was obtained from the Ghana Health Service Ethics Review Committee (approval 018/05/19). Informed consent was obtained from all participants after the purpose of the training was explained to them in English. Data from participants who opted out at any time from the study or incomplete data were not included in the final analysis. Privacy and confidentiality were ensured by completing the questionnaires individually, and the data were used following their consent. Participants were compensated (money for transportation and attendance and food such as breakfast, lunch, snacks, and take-home supper) for their participation in the training based on the number of days they participated. Our study adhered to the principles of the Declaration of Helsinki.

Results

BLS Training

A total of 229 health care professionals received BLS training, and immediately standardized questionnaires were administered for the assessment of knowledge and skills. The majority (152/229, 66.3%) were females. Most (107/229, 46.7%) were specialists or trainee specialists, with 48.1% (110/229) working in secondary health care facilities. Additionally, 26.2% (60/229) of the BLS training sessions were held in the Greater Accra region. The majority (171/229, 74.6%) demonstrated adequate knowledge and skills in BLS. We identified statistically significant associations between profession, facility, region, and BLS knowledge and skill scores (P<.05; Table 1).

Table 1. Sociodemographic characteristics of the respondents receiving basic life support training.

Variable	Total (N=229), n (%)	Basic life support knowledge score		*P* value
		Adequate (n=171), n (%)	Inadequate (n=58), n (%)	
Gender				.42
Female	152 (66.4)	111 (64.9)	41 (70.7)	
Male	77 (33.6)	60 (35.1)	17 (29.3)	
Profession				.03[a]
Physician assistant/nurse	86 (37.6)	57 (33.3)	29 (50)	
HO/MO/SMO[b]	36 (15.7)	32 (18.7)	4 (6.9)	
Specialist/trainee specialist	107 (46.7)	82 (48)	25 (43.1)	
Facility				.008[c]
Primary	64 (27.9)	57 (33.3)	7 (12.1)	
Secondary	110 (48)	76 (44.4)	34 (58.6)	
Tertiary	55 (24)	38 (22.2)	17 (29.3)	
Region				.03[a]
Ashanti	29 (12.7)	22 (12.9)	7 (12.1)	
Bono East	25 (10.9)	12 (7)	13 (22.4)	
Central	25 (10.9)	19 (11.1)	6 (10.3)	
Eastern	23 (10)	16 (9.4)	7 (12.1)	
Greater Accra	60 (26.2)	49 (28.7)	11 (19)	
Northern	32 (14)	23 (13.5)	9 (15.5)	
Volta/Oti	35 (15.3)	30 (17.5)	5 (8.6)	

[a]Differences were considered significant at P<.05.

[b]HO/MO/SMO: house officer/medical officer/senior medical officer.

[c]Differences were considered significant at P<.01.

Logistic regression analysis showed that respondents working in tertiary health care facilities were 80% less likely (adjusted odds ratio [AOR] 0.20, 95% CI 0.07-0.59; P=.003) to have adequate BLS knowledge and skills compared to those in primary health care facilities. Respondents in the Volta and Oti regions were 4.94 times more likely (AOR 4.94, 95% CI 1.17-20.80; P=.03) to possess adequate BLS knowledge and skills compared to those in the Bono East region (Table 2).

Table 2. Logistic regression analysis of the factors associated with respondents' basic life support knowledge and skill scores.

Variable	UOR[a] (95% CI)	*P* value	AOR[b] (95% CI)	*P* value
Profession				
Physician assistant/nurse	1.67 (0.89-3.14)	.11	1.45 (0.70-2.96)	.32
House officer/medical officer/senior medical officer	0.41 (0.13-1.27)	.12	0.47 (0.14-1.58)	.22
Specialist/trainee specialist	Ref[c]	Ref	Ref	Ref
Facility				
Primary	Ref	Ref	Ref	Ref
Secondary	1.00 (0.50-2.02)	>.99	0.97 (0.44-2.12)	.94
Tertiary	0.26 (0.10-0.71)	.009	0.20 (0.07-0.59)	.003[d]
Region				
Ashanti	1.91 (0.54-6.82)	.32	1.16 (0.30-4.49)	.83
Bono East	Ref	Ref	Ref	Ref
Central	1.90 (0.51-7.08)	.34	1.48 (0.37-5.97)	.58
Eastern	2.63 (0.72-9.61)	.15	1.29 (0.30-5.50)	.73
Greater Accra	1.35 (0.43-4.26)	.61	0.95 (0.27-3.36)	.93
Northern	2.35 (0.69-7.96)	.17	1.48 (0.38-5.82)	.58
Volta/Oti	6.50 (1.90-22.23)	.003	4.94 (1.17-20.80)	.03[e]

[a]UOR: unadjusted odds ratio.

[b]AOR: adjusted odds ratio.

[c]Ref: reference category.

[d]Significant at $P<.01$.

[e]Significant at $P<.05$.

ACLS Training

In total, 124 health care professionals received ACLS training, and standardized questionnaires were administered for the assessment of knowledge and skills. Most (73/124, 58.8%) were males. The majority (89/124, 71.7%) were doctors (house officer/medical officer/senior medical officer). More than half (64/124, 51.6%) worked in tertiary health care facilities. Most (55/124, 44.3%) of the ACLS trainings took place in the Greater Accra region. The majority (91/124, 73.4%) showed adequate knowledge and skills in ACLS. Statistically significant differences in ACLS knowledge and skills scores were observed based on gender, profession, and facility ($P<.05$; see Table 3).

Table 3. Sociodemographic characteristics of the respondents receiving advanced cardiac life support training.

Variable	Total (n=124), n (%)	Advanced cardiac life support knowledge score, n (%)		*P* value
		Adequate (n=91)	Inadequate (n=33)	
Gender				<.001[a]
Female	51 (41.1)	26 (28.6)	25 (75.8)	
Male	73 (58.9)	65 (71.4)	8 (24.2)	
Profession				.003[b]
Physician assistant/nurse	4 (3.2)	0 (0)	4 (12.1)	
House officer/medical officer/senior medical officer	89 (71.8)	68 (74.7)	21 (63.6)	
Specialist/trainee specialist	31 (25)	23 (25.3)	8 (24.2)	
Facility				.045[c]
Primary	20 (16.1)	12 (13.2)	8 (16.1)	
Secondary	40 (32.3)	26 (28.6)	40 (32.3)	
Tertiary	64 (51.6)	53 (58.2)	64 (51.6)	
Region				.34
Ashanti	12 (9.7)	10 (11)	2 (6.1)	
Bono East	11 (8.9)	8 (8.8)	3 (9.1)	
Central	11 (8.9)	5 (5.5)	6 (18.2)	
Eastern	11 (8.9)	9 (9.9)	2 (6.1)	
Greater Accra	55 (44.4)	39 (42.9)	16 (48.5)	
Northern	12 (9.7)	10 (11)	2 (6.1)	
Volta/Oti	12 (9.7)	10 (11)	2 (6.1)	

[a]Differences were considered significant at *P*<.001.

[b]Differences were considered significant at *P*<.01.

[c]Differences were considered significant at *P*<.05.

Compared to their female counterparts, males were 7.05 times more likely (AOR 7.05, 95% CI 2.69-18.46; *P*<.001) to have adequate knowledge and skills in ACLS (see Table 4).

Table 4. Logistic regression analysis of the factors associated with respondents' advanced cardiac life support knowledge and skill scores.

Variable	UOR[a] (95% CI)	*P* value	AOR[b] (95% CI)	*P* value
Gender				
Female	Ref[c]	Ref	Ref	Ref
Male	7.81 (3.12-19.54)	<.001	7.05 (2.69-18.46)	<.001[d]
Profession				
Physician assistant/nurse	__[e]	—	—	—
House officer/medical officer/senior medical officer	0.89 (0.35-2.28)	.80	0.66 (0.22-1.98)	.46
Specialist/trainee specialist	Ref	Ref	Ref	Ref
Facility				
Primary	3.21 (1.06-9.70)	.04	2.96 (0.79-11.13)	.11
Secondary	2.59 (1.04-6.50)	.04	2.25 (0.80-6.30)	.12
Tertiary	Ref	Ref	Ref	Ref

[a]UOR: unadjusted odds ratio.

[b]AOR: adjusted odds ratio.

[c]Ref: reference category.

[d]Significant at *P*<.001.

[e]Not applicable.

Discussion

Principal Findings

This study aims to evaluate the knowledge and skills of health care professionals immediately after training in Ghana on managing cardiac arrest through BLS and ACLS. We found that 75% of the health care professionals had adequate knowledge and skills in BLS, which was notably higher than that reported in other studies [2,9,11,12]. In similar training contexts, these percentages were 12% in Nepal [9] and 54% in Yemen [12]. Our results differ from studies in Afghanistan [2] and Pakistan [11], where 5% and 42% of respondents, respectively, had adequate BLS knowledge. The differences across these studies could be attributed to variations in research methods, the cutoff points used for determining knowledge levels, the effectiveness of the training programs, and whether assessments were conducted immediately after training. Additionally, prior BLS training may account for these variations, as Chaudhary et al [9] reported a strong correlation between knowledge scores and previous training.

This study reveals significant associations between BLS knowledge and skill scores and profession, facility, and region; however, no research has shown an association between the knowledge and skill scores of professions and facility type or region for comparison. We found that health care professionals working in tertiary health care facilities were less likely to have adequate knowledge and skills about BLS than those working in primary health care facilities. The reason could be that primary health care facilities in Ghana often provide basic emergency care as part of their primary services to communities; therefore, they receive more focused and frequent training on BLS techniques as part of their routine responsibilities compared to tertiary health care facilities, which prioritize specialized care for severe medical conditions, resulting in less emphasis on BLS training. The disparity in the findings could also be attributed to the assumption of prior knowledge and skills from practicing in a tertiary health facility.

Our study shows that health care professionals in facilities within the Volta and Oti regions possessed more adequate knowledge and skills in BLS compared to those in the Bono East region. This may be due to differences in the frequency of BLS training programs conducted in these areas. Facilities in the Volta and Oti regions might have more regular BLS training sessions than those in Bono East. Additionally, regions with better access to resources may facilitate more effective training for health care professionals, resulting in higher retention of BLS knowledge and skills; however, there is no evidence in the literature to support this justification.

Regarding ACLS, this study shows that 73% of the participants possessed adequate knowledge and skills, which contrasts with that reported in studies conducted in Ethiopia [1] and Spain [4]. Those studies showed that the majority, at 40% and 28%, respectively, had adequate knowledge of ACLS [1,4]. The difference may relate to variations in study methodologies, how knowledge was classified, and the previous training received on ACLS; none of these assessments were based on AHA-certified ACLS training. Adal and Emishaw [1] revealed that health care professionals who had previously received ACLS training were 5 times more knowledgeable than those who had not.

We found a significant difference between the ACLS knowledge and skill scores and gender. Compared with their female counterparts, male health care professionals were more likely to have adequate knowledge and skills in ACLS. This may be because our study had more male participants than female

counterparts. Another reason could be the differences in the types or frequency of the clinical experiences.

Strengths and Limitations

This study's strength is that it is the first in Ghana to evaluate health care professionals' knowledge and skills in managing cardiac arrest by using BLS and ACLS. It also included professionals from various backgrounds, facility types, and regions, thereby providing a thorough assessment across different parts of Ghana. In addition, our study created a standard for BLS and ACLS training and assessment that could be used in Ghana and other countries to support and monitor health care professionals' proficiency.

However, our study has some limitations. First, the evaluation of BLS and ACLS knowledge and skills was based on posttraining tests, which may not accurately reflect participants' real-world performance in emergencies and could introduce social desirability bias. Second, participants were recruited from only 8 regions and may not represent all the health care professionals in Ghana, potentially leading to selection bias. Third, there are no data on the specific deficiencies encountered.

Conclusion

Given an opportunity to learn and practice, health care professionals in Ghana attain adequate knowledge and skills in BLS and ACLS. BLS knowledge and skill scores were significantly associated with profession, facility, and region. Additionally, ACLS knowledge and skill scores were found to be associated with gender, profession, and facility. It is recommended that BLS and ACLS training be provided regularly to all health care professionals across all regions to improve their ability to respond effectively to medical emergencies.

Acknowledgments

The authors are very grateful to the participants of the cardiovascular disease training and who have contributed to the capacity building for cardiovascular disease management in Ghana. We are also grateful to the American Heart Association–certified instructors who delivered the training.

Data Availability

The data supporting the study's findings are available from the corresponding author upon request.

Funding

The Ghana Heart Initiative was implemented by the German International Cooperation (project 68.3025.1-001.00) in collaboration with the Ministry of Health and the Ghana Health Service, with funding from Bayer. The funder was not involved in the study design, data collection, analysis, interpretation, or writing of this manuscript.

Authors' Contributions

AD conceptualized the study and the initial draft of the manuscript. All the authors reviewed the manuscript and made constructive comments and recommendations. The authors approved the final version.

Conflicts of Interest

None declared.

References

1. Adal O, Emishaw S. Knowledge and attitude of healthcare workers toward advanced cardiac life support in Felege Hiwot Referral Hospital, Bahir Dar, Ethiopia, 2022. SAGE Open Med 2023;11:20503121221150101 [FREE Full text] [doi: 10.1177/20503121221150101] [Medline: 36685795]
2. Nemat A, Nedaie M, Essar M, Ashworth H, Aminpoor H, Sediqi A. Basic life support knowledge among healthcare providers in Afghanistan: a cross-sectional study of current competencies and areas for improvement. Ann Med Surg. 2023;5? 2023 Mar 16:8. [doi: 10.1097/ms9.0000000000000273] [Medline: 37113965]
3. Kumar R, Raja ., Badil ., Gulsher ., Ali J, Kumar A. Advanced cardiovascular life support performance: study on the effectiveness of training among nurses in a tertiary care hospital, Karachi. JPRI 2022 Feb 09:1-8. [doi: 10.9734/jpri/2022/v34i8b35478]
4. Pérez-Ibáñez R, Almodóvar-Fernández I, Gregori Roig P, Sánchez-Thevenet P. Level and determinants of advanced life support knowledge among nurses in Spain: a national cross-sectional study. SAGE Open Med 2025;13:20503121251326655 [FREE Full text] [doi: 10.1177/20503121251326655] [Medline: 40115866]
5. Doku A, Tuglo LS, Boima V, Agyekum F, Aovare P, Ali Abdulai M, et al. Prevalence of cardiovascular disease and risk factors in Ghana: a systematic review and meta-analysis. Glob Heart 2024;19(1):21 [FREE Full text] [doi: 10.5334/gh.1307] [Medline: 38404614]
6. Akumiah F, Yakubu A, Ahadzi D, Tuglo L, Mishra S, Mohapatra R. Cardiovascular care in Africa: cost crisis and the urgent need for contextual health service solutions. Glob Heart 2023 Aug 30:3. [doi: 10.5334/gh.1259] [Medline: 37664819]

7. Panchal A, Berg K, Kudenchuk P, Del RM, Hirsch K, Link M. 2018 American Heart Association Focused Update on Advanced Cardiovascular Life Support Use of Antiarrhythmic Drugs During and Immediately After Cardiac Arrest: An Update to the American Heart Association Guidelines for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care. Circulation 2018;American Heart Association focused update on advanced cardiovascular life support use of antiarrhythmic drugs during and immediately after cardiac arrest:a. [doi: 10.1161/cir.0000000000000613] [Medline: 30571262]
8. Panchal AR, Bartos JA, Cabañas JG, Donnino MW, Drennan IR, Hirsch KG, Adult BasicAdvanced Life Support Writing Group. Part 3: adult basic and advanced life support: 2020 American Heart Association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care. Circulation 2020 Oct 20;142(16_suppl_2):S366-S468 [FREE Full text] [doi: 10.1161/CIR.0000000000000916] [Medline: 33081529]
9. Chaudhary GP, Sah K, Malla J, Das N, Chaudhary S, Chaudhary I, et al. Knowledge regarding basic life support among health care workers of the hospital of Nepal. J Healthc Eng 2023;2023(9936114):9936114 [FREE Full text] [doi: 10.1155/2023/9936114] [Medline: 36644299]
10. Mahembe G, Mash R. The competency of nurses in basic life support at district hospitals in Cape Town, South Africa. S Afr Fam Pract (2004) 2025 May 29;67(1):e1-e8. [doi: 10.4102/safp.v67i1.6056] [Medline: 40459117]
11. Irfan B, Zahid I, Khan MS, Khan OAA, Zaidi S, Awan S, et al. Current state of knowledge of basic life support in health professionals of the largest city in Pakistan: a cross-sectional study. BMC Health Serv Res 2019 Nov 21;19(1):865 [FREE Full text] [doi: 10.1186/s12913-019-4676-y] [Medline: 31752855]
12. Alkubati SA, McClean C, Yu R, Albagawi B, Alsaqri SH, Alsabri M. Basic life support knowledge in a war-torn country: a survey of nurses in Yemen. BMC Nurs 2022 Jun 06;21(1):141 [FREE Full text] [doi: 10.1186/s12912-022-00923-0] [Medline: 35668520]
13. Doku AK, Tetteh J, Edzeame J, Peters RJG, Agyemang C, Otchi EH, et al. The Ghana heart initiative - a health system strengthening approach as index intervention model to solving Ghana's cardiovascular disease burden. Front Public Health 2024;12:1330708 [FREE Full text] [doi: 10.3389/fpubh.2024.1330708] [Medline: 38694980]
14. 2021 Population and housing census. Ghana Statistical Service. URL: https://statsghana.gov.gh/gssmain/fileUpload/pressrelease/2021%20PHC%20General%20Report%20Vol%203E_Economic%20Activity.pdf [accessed 2021-07-11]
15. Doku A, Tuglo LS, Chilunga F, Edzeame J, Peters RJ, Agyemang C. A multilevel and multicenter assessment of health care system capacity to manage cardiovascular diseases in Africa: a baseline study of the Ghana Heart Initiative. BMC Cardiovasc Disord 2023 Aug 24;23(1):421 [FREE Full text] [doi: 10.1186/s12872-023-03430-5] [Medline: 37620790]

Abbreviations

ACLS: advanced cardiac life support
AHA: American Heart Association
AOR: adjusted odds ratio
BLS: basic life support
CPR: cardiopulmonary resuscitation
ECC: emergency cardiovascular care

Edited by A Mavragani; submitted 05.04.25; peer-reviewed by GJ Noordergraaf, MS Shafi; comments to author 08.06.25; revised version received 28.06.25; accepted 16.09.25; published 31.12.25.

Please cite as:
Doku A, Tuglo LS, Osuoji C, Edzeame J, Broni M, Mainoo DD, Acquah AE, Honny K, Peters RJG, Agyemang C
Ghana Heart Initiative Training for Cardiac Arrest Management Among Health Care Professionals: Outcomes Evaluation Study (2019-2024)
JMIR Form Res 2025;9:e75536
URL: https://formative.jmir.org/2025/1/e75536
doi:10.2196/75536
PMID:

Predicting Ultra-High Risk Outcomes Using Linguistic and Acoustic Measures From High-Risk Social Challenge Recordings: mHealth Longitudinal Cohort Exploratory Study

Samuel Ming Xuan Tan[1], PhD; May Yen Lieu[1], BA; Jun Kai[2], BA; Zixu Yang[3], MSc; Luke KK[2], PhD; May O Lwin[4], PhD; Jimmy Lee[3], MD; Wilson Wen Bin Goh[1], PhD

[1]LKC School of Medicine, Nanyang Technological University, 59 Nanyang Drive, Experimental Medicine Building, Singapore, Singapore

[2]School of Humanities, Nanyang Technological University, Singapore, Singapore

[3]Institute of Mental Health, Singapore, Singapore

[4]WKW School of Communications, Nanyang Technological University, Singapore, Singapore

Corresponding Author:
Wilson Wen Bin Goh, PhD
LKC School of Medicine, Nanyang Technological University, 59 Nanyang Drive, Experimental Medicine Building, Singapore, Singapore

Abstract

Background: Early detection of individuals at ultra-high risk (UHR) for psychosis is critical for timely intervention and improving clinical outcomes. However, current UHR assessments, which rely heavily on psychometric tools, often suffer from low specificity. Speech-based machine learning prediction models can potentially be used to improve prognostic accuracy. However, existing studies often used long, open-ended speech tasks, which limit scalability. The High-Risk Social Challenge (HiSoC) is a short 45-second speech task designed to measure social functioning in individuals with UHR. If the HiSoC task is able to capture predictive signals, it may serve as an effective and scalable speech task for future prediction models.

Objective: The study aims to explore whether linguistic and acoustic features extracted from the HiSoC task are associated with UHR outcomes and if they are predictive of different UHR outcomes.

Methods: Audio recordings of HiSoC task responses were collected from 41 participants with UHR enrolled in the Longitudinal Youth at Risk Study. A total of 12 individuals converted to psychosis, 15 remitted from UHR status, and 14 maintained UHR status. The responses from the converted group were obtained within 12 months of psychosis onset, while the responses from the remitted and maintained groups were collected at baseline. Linguistic features analyzed included words per minute, articulation rate, dysfluency, and sequential coherence. Acoustic features comprised the mean and SD of fundamental frequency, the mean and SD of intensity, and HF500. Feature differential analysis was conducted via multivariate linear regression. Linear support vector machines were trained as outcome prediction models. Nested cross-validation was used to estimate the generalizability error. The models were principally evaluated on balanced accuracy (BA).

Results: The converted group exhibited lower words per minute (adjusted P=.02) and higher dysfluency (adjusted P=.004) compared to the remitted group. No significant differences were found in articulation rate, sequential coherence, or acoustic measures across the outcome groups. Two models outperformed random guess, namely the models using linguistic variables (BA 0.741, 95% CI 0.521-0.882) and linguistic and acoustic variables (BA 0.851, 95% CI 0.508-0.944).

Conclusions: Linguistic features extracted from a short speech task exhibit a measurable difference between the outcome groups. Our findings support the feasibility of using signals extracted from the HiSoC task recordings to predict remission in participants with UHR.

(JMIR Form Res 2025;9:e75960) doi:10.2196/75960

KEYWORDS

machine learning; mental health; outcome prediction; psychosis; speech data; ultra-high risk

Introduction

Psychosis is typically characterized by hallucinations without insight, delusions, and formal thought disorder [1]. Individuals who experience psychosis often experience a substantial decrease in their quality of life and might require long-term treatment with antipsychotic medication [2,3].

The accurate identification of individuals at heightened risk of developing psychosis is a key component in early intervention to improve clinical outcomes [4]. Many individuals who develop

psychosis often exhibit a prodromal phase during which subthreshold symptoms begin to manifest [5,6]. The identification of individuals in this prodromal phase is the basis of ultra-high risk (UHR) assessments such as the Structured Interview for Prodromal Syndrome and the Comprehensive Assessment of At-Risk Mental States [7,8]. However, these psychometric assessments often have high sensitivity but low specificity; that is, most individuals designated as UHR do not go on to develop psychosis [9]. Thus, there is substantial motivation to develop methods to supplement standard UHR assessments. Recently, an impressive range of prediction models has been developed using a variety of modalities, including biomolecular markers, clinical assessments, and linguistic and acoustic analyses [10-16].

Linguistic and acoustic analyses are particularly promising approaches since speech disturbances constitute some of the hallmarks of neurological disturbances and can be observed in most individuals with schizophrenia [17,18]. Deficits such as poverty of speech, greater dysfluency, reduced coherence, derailment, and tangentiality have been consistently reported over the years and form what is now commonly known as "schizophrenia speech" [19-22]. The presence of such deficits is often correlated with limited functioning [23]. Some of these deficits can often be observed at early stages of disease progression, including individuals with UHR [24,25]. Additionally, UHR individuals displaying greater deficits in verbal fluency and coherence are more likely to transition to psychosis [26-28]. There are also linguistic differences between individuals with early stages of schizophrenia and individuals with established schizophrenia, suggesting that the deficits can vary across the disease progression [29]. These varied speech and linguistic deficits are consistently observed across different languages and cultures, including Japanese-, Chinese-, and Portuguese-speaking individuals with UHR [30-34].

Various studies have attempted to combine natural language processing (NLP) methods and machine learning to predict UHR outcomes. One example [10] used open-ended narrative interviews of approximately 1-hour duration from 34 individuals with UHR, with 5 converting to psychosis, to train models using semantic coherence and speech complexity, achieving 100% accuracy in predicting psychosis onset. A more recent work [12] involved developing a predictive model using speech data from the Caplan "story game" along with linguistic markers, such as reduced semantic coherence, increased variance in coherence, and decreased use of possessive pronouns [35]. The study used 93 participants with UHR recruited from 2 sites and achieved 83% accuracy in predicting psychosis onset. Finally, current literature also reported that individuals with UHR with lower connectedness at baseline are more likely to develop affective disorders [33].

These studies suggest potential for the use of NLP methods and machine learning on speech recordings to predict UHR outcomes. Automatic speech recognition (ASR) technologies have advanced substantially in recent years. While much work remains in terms of ensuring the reliable performance of ASR models in real-world applications and in individuals with dysfluent speech [36,37], the general trends are promising—with some models achieving over 90% accuracy in benchmark tests [38,39]. Recent automated speech analysis pipelines have also found some success in predicting an individual's depression, anxiety, and suicidal ideation level as assessed by self-reported questionnaires [40-42]. With the eventual development of more accurate ASR models, it is conceivable that psychosis risk screens based on automated voice and speech analysis can be developed in the near future. For such screens, long and open-ended speech tasks such as those used in [10,12] might not be scalable as they usually require a lengthy involvement and trained personnel to administer the task. As such, we believe that it is appropriate to explore the predictive potential of speech data extracted from shorter speech tasks. Such findings will be useful in identifying potential tasks that can be more readily used in future automated screens.

The High-Risk Social Challenge (HiSoC) task is designed to assess social functioning in individuals with UHR [43,44]. In the HiSoC task, participants are tasked with providing a 45-second response to a scenario, such as an audition for a competition or a job interview, with minimal preparation. Participant responses are video-recorded and scored on 16 items by trained assessors on a 5-point Likert scale. We previously demonstrated that the HiSoC task can effectively discriminate between individuals with UHR and healthy controls [43,45]. Several properties of the HiSoC task make it a particularly promising source of prognostic information. First, the HiSoC task can be administered quickly, requiring only 10 seconds of preparation and 45 seconds for execution (approximately 1 min total). Second, it is designed to evaluate social functioning, which has been consistently reported to be a strong predictor of clinical outcome [46-48]. Third, the HiSoC task collects video recordings from which audio recordings can be extracted. The speech contents of the recordings can be transcribed, and the acoustic properties of the speech are analyzed to generate a significant amount of data points for research and potential prognostic purposes. Fourth, the HiSoC task only requires a medium through which the prompt can be transmitted and a device to capture a video of the response; both of which can be done using a smartphone. These properties suggest the HiSoC is a task that is potentially suitable for future screens, and there is potential for the screen to be completely remote and automated.

In this study, we perform an exploratory study on the feasibility of using linguistic and acoustic features extracted from HiSoC task recordings to predict outcomes in UHR. The data used in this study were collected as part of the LYRIKS (Longitudinal Youth at Risk Study) [49], an Asian UHR cohort. We examined 2 prediction outcomes, conversion and remission. While the prediction of conversion is of obvious clinical importance, the ability to accurately predict remission is also clinically important, as it allows for individuals who are likely to remit to be assigned to a lower risk group. More intensive intervention can then be directed toward those who are at a higher risk of conversion and maintaining UHR status. Indeed, individuals who maintain UHR status often still experience reduced functioning and long-term attenuated psychotic symptoms[50].

Methods

Participants

Participants were recruited as part of the LYRIKS [49]. The LYRIKS is a longitudinal cohort observation study conducted between 2008 and 2010. A total of 2368 individuals were assessed for eligibility, the Comprehensive Assessment of At-Risk Mental States was performed for 926 individuals, and 667 were accepted into the study. The 667 participants consist of 173 participants with UHR and 494 control participants aged between 14 and 29 years. The participants were monitored over a 2-year period between 2008 and 2010. Of the 173 participants with UHR, 17 converted to psychosis (approximately 10% conversion rate). Participants who converted were removed from the study following the collection of the final data point.

Participants assessed to have converted to psychosis were excluded from the study following final data collection. Participants in the LYRIKS were recruited from a mixture of help-seeking and non–help-seeking individuals. Outreach and recruitment strategies are detailed in [51]. All assessments were performed at the same center (Institute of Mental Health, Singapore). The inclusion criteria for the study include (1) aged between 14 and 29 years and (2) English-speaking. Exclusion criteria include (1) having a past or current history of psychosis or intellectual disability, (2) currently using illicit substances, (3) taking antipsychotics or mood stabilizers, (4) having medical causes associated with their psychosis, and (5) contraindications for magnetic resonance imaging. None of the participants were exposed to antipsychotics, mood stabilizers, or illicit substances including cannabis.

Study participants were selected based on the availability of HiSoC recording data, which were collected at 12-month intervals (mo 0, mo 12, and mo 24). Of the 173 participants with UHR, 50 remitted from UHR status within the first 12 months of the study. Among the 17 participants with UHR who transitioned to psychosis, HiSoC task recordings from within the 12 months prior to conversion were available for 12 participants. All 12 recordings were included to form the Converted outcome group. A total of 32 UHR participants did not convert to psychosis but continued to meet the criteria for UHR throughout the duration of the study. HiSoC task recordings from month 0 are available for 14 of them. All 14 recordings were selected to form the maintained outcome group.

HiSoC task recordings from month 0 were available for 28 participants who remitted, and 15 were randomly selected to form the Remitted outcome group. This undersampling was performed to keep the number of individuals in each outcome group proportionally similar to avoid class imbalance issues during the training of predictive modeling classifiers.

HiSoC Task

Speech recordings used in this study were recorded as part of the HiSoC [43]. Participants were presented with a scenario where they are taking part in a “most interesting person in Singapore” competition, whereby “The winner will be selected based on a 45-second video about themselves.” The participants were given 10 seconds to prepare a response before video recording commenced. The video-recorded response was assessed by 2 trained raters on 16 items each on a 5-point Likert scale. The 16 items can be grouped into 5 domains: affect, social-interpersonal, behavior, and language [44]. The HiSoC task generates a video recording of the participant performing the task, along with the raters’ scoring. All HiSoC tasks were performed in the same study center and recorded using a Sony Handycam DCR SR47 camcorder.

Covariates

Various covariates were assessed to ensure that the outcome groups do not significantly differ in terms of symptom severity, anxiety, cognition, depression, and education levels. Symptom severity was measured using the Positive and Negative Syndrome Scale (PANSS), which is a clinical assessment of the severity of positive and negative symptoms in individuals with psychosis and UHR [52]. Anxiety was assessed using the Beck Anxiety Inventory (BAI) score, which is the total score across the 21 items of the BAI [53]. Cognitive performance was measured using the Brief Assessment of Cognition in Schizophrenia (BACS), which is an instrument that specifically assesses the aspects of cognition impaired and correlated with clinical outcomes in individuals with schizophrenia [54]. Aspects assessed by the BACS include verbal memory, disorganized speech, token motor task (TMT), verbal fluency, symbol coding, and the Tower of London. The presence of depressive disorder was assessed by whether the individual had an active diagnosis of a depressive disorder [1]. Education level was assessed by 2 measures, namely whether the participant undertook the Primary School Leaving Examination (PSLE) later than expected and whether they have a low education level relative to age. The PSLE is a mandatory national examination taken by all school children at 12 years of age in Singapore. We defined an individual to have late PSLE if they undertook the PSLE after the age of 13 years. Individuals were indicated as having low education relative to age if they had not attained or were currently undergoing postsecondary education by the age of 18 years.

Transcription

To maximize transcription accuracy, we used manual transcription by 2 independent transcribers (MYL and JK) trained in conversation analysis and transcription methodologies. The transcribers were blinded to the outcome group of the individuals in the recording. These transcribers were not trained in rating the HiSoC task. All identifiable information was removed from transcripts. Transcriber 1 completed all 41 recordings, while transcriber 2 transcribed 12 randomly selected recordings (4 from each outcome group). Consistency between the 2 transcribers was assessed using the Pearson correlation. VLC media player (VideoLAN) was used to extract audio files from the video recording [55]. Speech was performed using PRAAT (version 6.3.15; Boersma and Weenink) [56]. The transcription key used can be found in Table S1 in Multimedia Appendix 1.

The spectrogram was used to support the identification of silent segments, pitch, and intensity variations. Timestamped annotation and transcripts from PRAAT were exported as

textgrid files into Python (Python Software Foundation) for feature extraction.

Linguistic Variables

The following linguistic variables were extracted from the recordings:

1. Words per minute (WPM): the average number of words spoken by participants within 1 minute. However, since the duration of the HiSoC task is fixed at 45 seconds, our version of WPM is determined by multiplying the total number of words spoken during the task by 0.75.
2. Articulation rate (AR)*:* speed of speech production. It is determined by dividing the total word count by the actual speech duration, excluding pauses [57].
3. Dysfluency: the ratio of short or medium pauses, along with the number of interjections, to the total word count in a text. Short pauses are defined as those lasting less than 0.3 seconds, while medium pauses range between 0.3 and 0.7 seconds. Interjections are identified using spaCy's Part-of-Speech tagging, made available via the "en_core_web_lg" model [58].
4. Sequential coherence (SC)*:* connectedness and similarity between adjacent words. SC is effective in differentiating individuals with schizophrenia from healthy controls and in performing derailment detection [59,60]. Using Word2Vec embeddings from the spaCy en_core_web_lg model, SC is calculated as the mean Word2Vec similarity between adjacent words across the text [58,60]. A moving average with a window of size 5 was used. SC was computed using Word2Vec rather than distribution methods such as latent semantic analysis (LSA) and Latent Dirichlet Allocation, as distributed methods such as Word2Vec were reported to have better performance and more closely match human ratings [61,62].

All linguistic features and their abbreviations are listed in Table 1.

Table . Name and abbreviation of linguistic and acoustic features.

Type and variable name	Variable abbreviation
Linguistic	
Words per minute	WPM
Articulation rate	AR
Dysfluency	Dysfluency
Sequential coherence	SC
Acoustic	
F0 mean	F0_m
F0 SD	F0_sd
Intensity mean	Int_m
Intensity SD	Int_sd
HF500	HF500

Acoustic Variables

Intensity (loudness), fundamental frequency F0 (pitch), and spectral energy were extracted from audio recordings and used to derive the following acoustic variables:

1. Fundamental frequency (F0): the rate at which the vocal fold vibrates during speech. Fundamental frequency conveys key elements about the speaker's identity (different F0 across vowels), sex (lower in males), and emotion (higher and lower F0 when happy and sad, respectively) [63,64]. The *mean fundamental frequency (F0_m*) and *F0 standard deviation (F0_sd*) were extracted from each recording using PRAAT [56]. A high-pass filter at 140 Hz for female participants and 75 Hz for male participants, along with a low-pass filter of 300 Hz for both sexes, was applied.
2. Intensity: the loudness of the voice measured in decibels. We calculated the *mean intensity (Int_m*) and *intensity standard deviation (int_sd*) of the intensity values obtained from PRAAT [56]. These measures allow us to examine whether the different outcome groups exhibit differences in loudness and variations in loudness. Readings below 10 dB were omitted to reduce the effect of ambient sound on the measures.
3. HF500*:* the relative proportion of high-frequency acoustic energy (>500 Hz) to low-frequency acoustic energy (<500 Hz) in the spectrum. This measure has been reported to be a viable measurement of emotional states in voices [65].

All acoustic features and their abbreviations are listed in Table 1.

Data Processing and Statistical Analysis

Data processing and statistical analysis were conducted in the Python version 3.10 programming environment. The data were standardized prior to statistical testing and predictive modeling. Statistical significance between the outcome groups across covariates was assessed using ANOVA for continuous variables and the chi-square test for binary variables.

Linear regression models were constructed for each linguistic and acoustic feature. To allow for assessments on whether differences in linguistic and acoustic features are associated with depression diagnosis (DD), sex, cognition (BACS), or

anxiety (BAI), these covariates are included in the model along with the outcome group (outcome):

y~DD+sex+BACS+BAI+outcome

To examine pairwise differences between the outcome groups, we performed pairwise *t* tests on the outcome groups. Regression analyses were performed using the statsmodels 0.14.4 Python package. Multiple test correction was performed using the Benjamini-Hochberg procedure [66].

Outcome Prediction Modeling

Logistic regression and support vector machine (SVM) with a linear kernel are 2 commonly used machine learning models [67]. Mathematically, they are related and tend to perform comparably across most tasks [68]. However, there are some studies suggesting that the SVM performs slightly better in imbalanced datasets [69]. Since our predictive modeling task involves class imbalance, we opted to use SVMs in our study. We used linear SVM with balanced class weights from the *scikit-learn* Python package [70]. Given a dataset with N samples and K classes, the balanced class weight wi for class i is implemented as:

wi=NKni

where ni is the number of samples in class i.

To perform robust model training and evaluation, we used a nested cross-validation setup. This approach leverages an outer leave-one-out cross-validation loop for performance assessment while relying on an inner stratified 5-fold cross-validation loop for hyperparameter tuning. We selected the best-performing model from the inner loop and passed it to the hold-out test sample in the outer loop. Model output consists of the predicted class label.

We repeated the machine learning training process on 5 combinations of features: HiSoC verbal features only (HiSoC_ve), all HiSoC features (HiSoC_all), linguistic features (linguistic), acoustic features (acoustic), and linguistic and acoustic features (linguistic_acoustic). HiSoC_vs features consist of the items with a strong emphasis on participants' voice: verbal expression, clear communication, fluency of speech, and social anxiety. HiSoC_all consists of all 15 HiSoC items. linguistic_acoustic consists of all linguistic and acoustic features.

Given there are 3 outcome groups, one-vs-all classification was used to transform the task into a binary classification task. Model performances on 2 tasks were examined: predicting conversion outcome in the next 12 months (converted-vs-all) and predicting remission outcome in the next 12 months (remitted-vs-all). The converted-vs-all task consists of 12 converted individuals as the positive class and 29 nonconversion (15 remitted+14 maintained) individuals as the negative class. The remitted-vs-all task consists of 15 remitted individuals as the positive class and 26 nonremitted (12 converted+14 maintained) individuals as the negative class.

Model Evaluation

Model performance was assessed using balanced accuracy (BA), defined as:

BA=TPR+TNR2

where TPR and TNR are the true positive rate and true negative rate, respectively. 95% CIs for BA were constructed based on 1000 bootstrap resamples. Estimates of generalizability error were obtained from the outer fold of the nested cross-validation. 95% CIs are denoted in brackets in the "Results" section.

Common methods to assess overall model performance when significant data imbalances are present include the BA, the Matthew correlation coefficient (MCC), and the precision-recall curve. The precision-recall curve is not suitable for this study as it requires decision probabilities, and decision probabilities in the SVM in *scikit-learn* are derived via Platt scaling, which is a computationally intensive process that will be further compounded by the bootstrapping procedure [70,71]. We chose BA over MCC as it is often impossible to compare MCC of models trained on different datasets—a process necessary to facilitate future validation [72]. In an imbalanced dataset, classifying all samples to the majority class will give a BA of 0.5, which is equivalent to the expected BA of a random guess in a balanced dataset. We define a model performance to be statistically significant if it outperforms a random guess; that is, the lower bound of 95% CI for BA is >0.5.

Ethical Considerations

Ethical approval for the LYRIKS was provided by the National Healthcare Group's Domain Specific Review Board (approval: 2009/00167). After a complete description of the study was provided to the participants, written informed consent was obtained. Participants have the ability to opt out of any assessment or terminate participation at any time. Participants were compensated after each visit. All data used were deidentified prior to any analysis. Secondary analyses such as those performed in this study are fully covered under existing ethical approvals and written informed consent from the participants. All researchers involved were required to sign confidentiality and data protection agreements prior to access to the data.

Results

Demographics

Across the outcome groups, no significant differences in age, sex (proportion of female participants), PANSS, education (late PSLE and low education relative to age), and BAI scores were observed. Statistically significant differences in BACS TMT across the outcome groups were observed ($F_{2,22}$=5.214, P=.02; Table 2). The Tukey test revealed that the remitted group exhibited a significantly higher score for BACS_TMT than the Converted group (Table S2 in Multimedia Appendix 1), suggesting that the converted group has much lower motor speed than the remitted group.

Table . Participant demographics.

Characteristic	Remitted	Maintained	Converted	ANOVA (*P* value)	Chi-square test (*P* value)
Age (y), mean (SD)	22.1 (2.90)	20.6 (4.16)	20.9 (3.75)	.51	—[a]
Sex, n (%)				—	.68
Female	6 (40)	4 (28.6)	3 (25)		
Male	9 (60)	10 (71.4)	9 (75)		
PANSS[b], mean (SD)					
PANSS +	9.9 (2.99)	10.6 (2.56)	10.7 (2.87)	.72	—
PANSS –	10.6 (4.29)	12.1 (4.37)	12.6 (4.01)	.44	—
Education attainment, n (%)					
Late PSLE[c]	0 (0)	1 (7.14)	0 (0)	—	.38
Low education level relative to age	1 (6.67)	1 (7.14)	2 (16.7)	—	.64
BACS[d], mean (SD)					
VM[e]	43.1 (7.18)	45.1 (11.64)	43.2 (9.14)	.82	—
DS[f]	21.0 (4.07)	20.4 (3.89)	18.3 (4.01)	.22	—
TMT[g]	76.3 (8.17)	70.1 (12.09)	62.0 (13.86)	.02	—
VF[h]	47.3 (12.75)	41.8 (10.82)	37.4 (11.17)	.10	—
SC[i]	58.3 (10.48)	58.1 (9.48)	52.6 (16.77)	.42	—
TOL[j]	18.0 (1.69)	18.7 (2.20)	16.9 (3.40)	.19	—
Anxiety and depression					
BAI[k] score, mean (SD)	17.3 (13.15)	16.9 (11.41)	21.3 (15.13)	.67	—
DD[l], n (%)	4 (26.7)	3 (21.4)	4 (33.3)	—	.79

[a]Not applicable.
[b]PANSS: Positive and Negative Syndrome Scale.
[c]PSLE: Primary School Leaving Examination.
[d]BACS: Brief Assessment of Cognition in Schizophrenia.
[e]VM: verbal memory.
[f]DS: disorganized speech
[g]TMT: token motor task.
[h]VF: verbal fluency.
[i]SC: symbol coding.
[j]TOL: tower of London.
[k]BAI: Beck Anxiety Inventory.
[l]DD: depression diagnosis.

Linguistic Measures

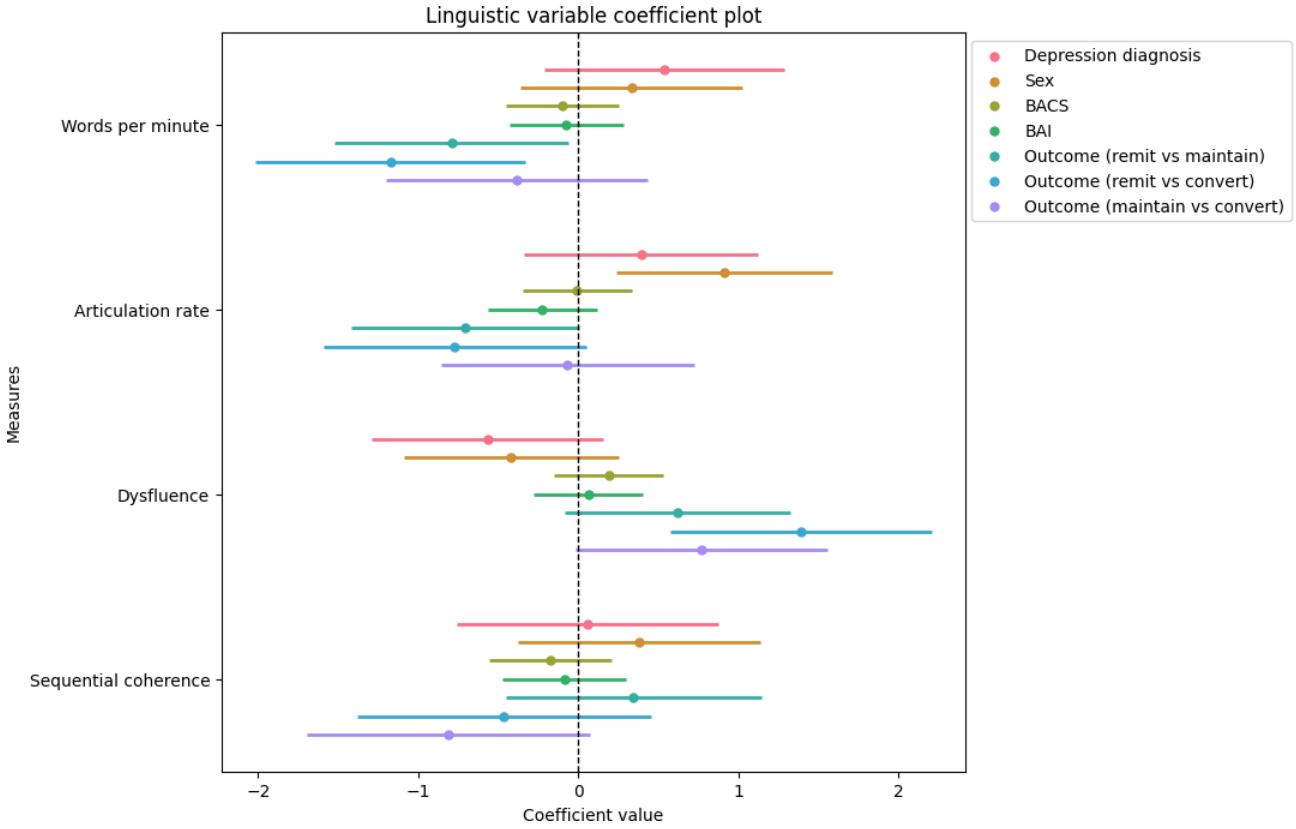
WPM, AR, dysfluency, and SC measures were consistent between transcribers (R^2=0.993, 0.993, 0.929, and 0.868 for WPM, AR, dysfluency, and SC, respectively; Figure S1A-D in Multimedia Appendix 1), indicating that the transcription and linguistic measures are consistent across transcribers.

WPM was lower in the maintained group relative to the remitted group (β=−0.79, 95% CI −1.52 to 0.06; *P*=.04); however, this difference was no longer significant following FDR correction (adjusted *P*=.05). Similarly, WPM was lower in the converted group compared to the remitted group (β=−1.17, 95% CI −2.02 to −0.33; *P*=.008). This result remained significant following FDR correction (adjusted *P*=.02). Since AR was not observed to significantly differ between the outcome groups, this reduction in WPM suggests the converted group spoke at a similar speed as the remitted group but spoke significantly fewer words.

We also observed that the converted group exhibits significantly higher dysfluency relative to the remitted group (β=1.39, 95%

CI 0.58-2.21; *P*=.001), surviving FDR correction (adjusted *P*=.004; Figure 1), suggesting that the speech of the converted group has significantly more interjections and pauses.

Figure 1. Coefficient plot of each covariate for each linguistic measure (outcome). The coefficients of models fitted to words per minute, articulation rate, dysfluency, and sequential coherence are shown. Each point represents the estimated coefficient for a given predictor-response pair, with horizontal lines indicating the 95% CIs. To facilitate interpretation, we presented coefficients of the outcome group contrasts rather than the coefficients of the outcome group covariates. The covariate is statistically significant if the 95% CI does not intersect 0. BACS: Brief Assessment of Cognition in Schizophrenia; BAI: Beck Anxiety Inventory.

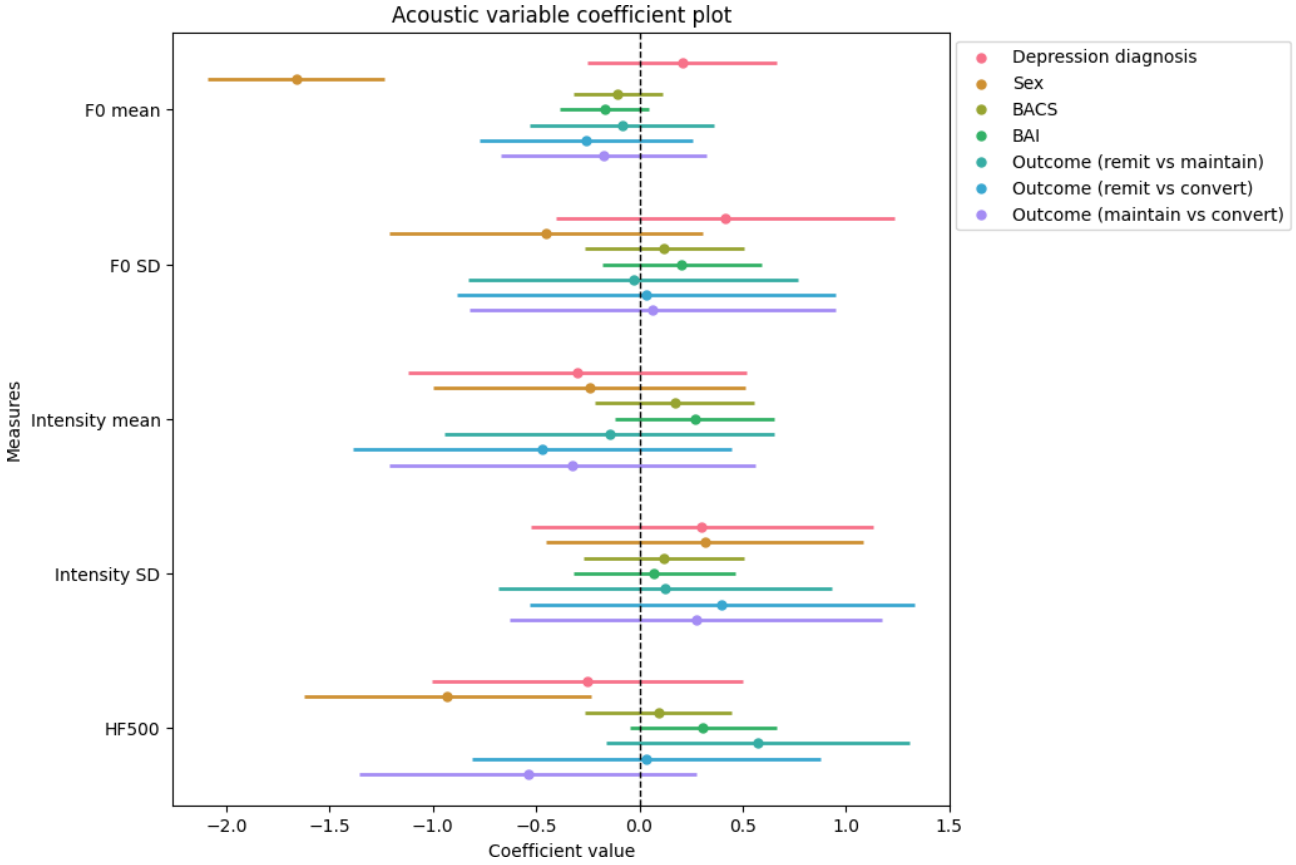


A full table of all coefficients and the associated statistics can be found in Tables S3-S4 in Multimedia Appendix 1.

Acoustic Measures

We did not observe any significant differences between the outcome groups across all 5 acoustic measures. We observed sex differences in F0 mean and HF500, with male participants exhibiting lower F0 (β=−1.66, 95% CI −2.09 to −1.24; *P*<.001) and lower HF500 (β=−1.66, 95% CI −2.09 to −1.24; *P*<.001) than female participants (Figure 2). These observations indicate a lower pitch in male participants and a brighter voice quality in female participants. These are expected differences.

Figure 2. Coefficient plot of each covariate for each acoustic measure (outcome). The coefficients of models fitted to F0_m, F0_sd, Int_m, Int_sd, and HF500 are shown. Coefficient plot of the acoustic measures for acoustic features. Each point represents the estimated coefficient for a given predictor-response pair, with horizontal lines indicating the 95% CIs. To facilitate interpretation, we presented coefficients of the outcome group contrasts rather than the coefficients of the outcome group covariate. The covariate is statistically significant if the 95% CIs do not intersect 0. BACS: Brief Assessment of Cognition in Schizophrenia; BAI: Beck Anxiety Inventory.



A full table of all coefficients and the associated statistics can be found in Tables S5-S6 in Multimedia Appendix 1.

Outcome Prediction

We examined the performance models trained using HiSoC_all, HiSoC_ve, linguistic, acoustic, and linguistic+acoustic features set across the converted-vs-all and remitted-vs-all tasks.

In the converted-vs-All task, the acoustic model demonstrated the highest BA (BA=0.595, 95% CI 0.282-0.764), followed by the linguistic model (BA=0.570, 95% CI 0.339-0.815) and HiSoC_ve (BA=0.480, 95% CI 0.203-0.774). The HiSoC_all model (BA=0.470, 95% CI 0.310-0.778) and the linguistic+acoustic model (BA=0.529, 95% CI 0.246-0.798) achieved the lowest BAs in this task. However, none of the model performances outperformed a random guess as the lower bounds of the 95% CI of BA were <0.5.

In the remitted-vs-all task, the linguistic+acoustic model achieved the highest balanced accuracy (BA=0.851, 95% CI 0.508-0.944), followed by HiSoC_all (BA=0.760, 95% CI 0.382-0.9), linguistic (BA=0.741, 95% CI 0.521-0.882), and HiSoC_ve (BA=0.645, 95% CI 0.405-0.813). The acoustic model demonstrated the lowest balanced accuracy (BA=0.574, 95% CI 0.325-0.798) in this task. The performances of the linguistic+acoustic model and the linguistic model both outperform a random guess. However, there is substantial overlap between the 95% CI of the 2 models, which means that we cannot determine if there are any meaningful differences in performance between the 2 models.

Regularization parameters and model coefficients are provided in Tables S7-S9 in Multimedia Appendix 1. Specificity and sensitivity of the models are provided in Table S10 in Multimedia Appendix 1.

Discussion

Principal Findings

In this study, we explore the outcome prediction potential of linguistic and acoustic features extracted from the HiSoC task. Our findings suggest that linguistic and acoustic features extracted from the HiSoC task contain signals that can potentially differentiate between the outcome groups; most notably, the converted group exhibits lower WPM and higher dysfluency compared to the remitted group. In our prediction task, our linguistic and linguistic+acoustic models achieve good performance (BA=0.741 and 0.851, respectively) and outperformed random guess in the remitted-vs-all task. These findings are promising and support further studies around the use of short speech tasks such as the HiSoC for outcome prediction.

Regression analysis revealed the converted group exhibited lower WPM and higher dysfluency relative to the remitted group. The decrease in WPM in the converted group is indicative

of the poverty of content. This reduction is consistent with reduced speech time in individuals with schizophrenia compared to healthy controls [73]. Measures of poverty of speech, both via expert evaluation and NLP methods, have been shown to be predictive of psychosis onset in individuals with UHR [16,74,75]. The increase in dysfluency in the converted group compared to the remitted group is consistent with reports of individuals with UHR who convert to psychosis displaying greater dysfluency compared to those who do not [26,27]. Additionally, greater dysfluency is correlated with increased negative symptom severity, which is in turn correlated with an increased risk of psychosis onset [76,77].

We did not observe any statistically significant differences in SC. This is despite a reduction in semantic coherence being a key predictor of conversion outcome in prior studies [10,12]. We hypothesize two reasons for this difference: (1) this could be due to the length of the HiSoC task being too short to effectively collect sufficient speech output for semantic coherence to be accurately measured. (2) Semantic coherence in this study is measured as SC, which is the average Word2Vec similarity between adjacent words, whereas LSA was used in prior studies [10,12]. The SC method was chosen as Word2Vec had been shown to outperform LSA and is more consistent with human raters than distributional methods such as latent Dirichlet allocation and LSA [61,62]. It is possible that LSA is superior to Word2Vec in this application.

We also did not observe any statistically significant differences between the outcome groups in any of the acoustic measures assessed (F0 mean, F0 SD, intensity mean, intensity SD, and HF500) across the 3 outcome groups. This is despite monotonous speech being a common feature of schizophrenia speech [19]. Meta-analyses of voice patterns in schizophrenia have found that the effect sizes of reduced pitch variability are inconsistent across studies [73], suggesting that, despite monotonous speech being a common feature of schizophrenia speech, reduced pitch variability is not always observed. This could be due to the inherent heterogeneity in the manifestation of speech and language disturbances as well as the nature of the task used to generate the response [78]. The lower F0 mean and HF500 observed in male participants are expected sex differences.

In our prediction tasks, only the linguistic+acoustic model and the linguistic model in the remitted-vs-all task were able to outperform a random guess. This has 2 key implications. First, the primary purpose of this study is to explore the predictive potential of short speech tasks such as the HiSoC. With this result, we found evidence suggesting that linguistic and acoustic features extracted from the HiSoC task can capture speech features that are predictive of remission. Second, none of the models in the converted-vs-all task achieved a performance that is statistically significant, suggesting that the linguistic and acoustic features were able to predict remission but not conversion. Together with the lack of any statistical difference between the converted and maintained groups, it is suggested that the speech patterns of the maintained group do not differ significantly from the converted group within the HiSoC task. If this finding is generalizable, it suggests that the speech patterns of individuals who convert to psychosis and individuals who maintain UHR status are largely similar. Consequently, efforts to predict conversion to psychosis using speech patterns will always be complicated by difficulties in differentiating between individuals who converted and individuals who maintained. A recent study has found that language disturbances are a strong predictor of response to clinical interventions; individuals with UHR with lower levels of language disturbances exhibit greater improvement in both symptom severity and functioning over time [50]. It is possible that speech and language disturbances more accurately reflect individual capacity for improvement rather than eventual clinical outcome. With these considerations, predicting remission from UHR status might be a more feasible direction than predicting conversion to psychosis. The ability to identify individuals likely to remit still has tremendous use as it allows for greater focus to be placed on those not likely to remit, allowing limited resources to be distributed to those who need them the most.

While our findings indicate that signals extracted from the HiSoC task can feasibly be used to predict remission, it must be reiterated that the study is intended to be exploratory and that any findings are exploratory and limited by the small sample size. Even so, signals are still strong enough to be detected. Future validation studies with larger independent datasets are necessary to validate both the findings and model generalizability before clinical or screening implications can reasonably be considered.

This study examines predictive potential involving speech data extracted from the HiSoC task. However, while there are several tasks designed to elicit speech in mental health, there is little consistency in the tasks used. For example, tasks used in recently published automated speech analysis pipelines include reading from selected passages [40], semistructured speech tasks such as “Describe how you are feeling at the moment and how your nights’ sleep have been lately” [42], and talking to research nurses [41]. A comparative study using a variety of speech tasks should be performed to examine whether the outcome group differences are consistent across different tasks, and if there is an optimal task for outcome prediction.

While ASR promises scalability that can potentially unlock fast and efficient automated speech-based risk screens, current ASR models tend to exhibit higher error rates in dysfluent speech [36,37]. This might be particularly problematic in psychosis risk screens, where dysfluency is a feature of schizophrenia speech. ASR technologies will likely need to reach a sufficiently reliable and consistent accuracy before an automated psychosis risk screen can achieve sufficient reliability.

Strengths

To our knowledge, this is the first study diving into the predictive potential of linguistic and acoustic features extracted from audio recordings of the HiSoC task. The recordings used in this study are significantly shorter and more scalable than those in comparable studies [10,12]. While significant validation work remains, we showed that features from the HiSoC task contain statistically significant differences between the outcome groups and that extracted linguistic and acoustic features can be used to predict remission.

Our findings suggest that further exploration into the predictive use of short speech tasks such as the HiSoC in speech analysis is warranted. We expect that this study will be one of the first of many that explore or validate the predictive use of various short speech tasks to facilitate future speech−based automated risk screening tools.

Limitations

First, although convenient, the short duration of the HiSoC task can potentially lead to data that are less representative of the individual's speech pattern. As described previously, this might explain the lack of differences in SC between the outcome groups. Additional studies comparing longer open-ended speech tasks and shorter tasks like the HiSoC will be necessary to assess whether shorter tasks sufficiently capture the individual's speech patterns. Second, our sample sizes are limited by the undersampling performed to keep the number of individuals in each outcome group relatively balanced to minimize class imbalance issues. This meant that our sample size would be limited by the number of participants who converted to psychosis even when more data from individuals who remitted or maintained were available. A small sample size leads to lower statistical power of our regression analysis, which means that there might be differences between the outcome groups that were not detected due to the low statistical power of the test. The large 95% CIs for balanced accuracy in our models are likely a consequence of the small sample size, as the performance of the model can fluctuate significantly depending on the bootstrap resample. A small sample size can also lead to the creation of biased models that do not generalize well. However, the purpose of this study is to explore the potential of developing outcome prediction models using features extracted from the HiSoC task audio recordings and not to develop a definitive model. Third, we lack an independent validation dataset. This limits our ability to accurately estimate generalizability error. It is possible that any class separation within the feature space used in this study is unique to this dataset. A follow-up study using the same feature sets and methods on a comparable dataset is necessary to validate both the regression analysis findings and the model performances.

Acknowledgments

The authors declare the use of generative artificial intelligence (GAI) in the research and writing process. According to GAIDeT (Generative AI Delegation Taxonomy; 2025), the following tasks were delegated to GAI tools under full human supervision: reformatting (formatting of numerical values, *P* values, β, 95% CI, and so forth from tabular to text structure). The GAI tool used was ChatGPT 4.0. The responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Funding

WWBG acknowledges support from Ministry of Education Tier 1 (RS08/21 and RT11/21) awards. This research was supported by the National Medical Research Council, Singapore, under its Population Health Research Grant scheme (project PHRGOC24jul-0026). The Longitudinal Youth at Risk Study was supported by the National Research Foundation Singapore under the National Medical Research Council Translational and Clinical Research Flagship Programme (grant NMRC/TCR/003/2008). JL received funding support from the Singapore Ministry of Health's National Medical Research Council (grant MOH-CSAINV17nov-0004).

Data Availability

The data used in this study are not publicly available due to ethical and legal requirements. However, researchers who wish to access or investigate the data for valid scientific purposes may contact the corresponding author. All data sharing requests will be evaluated on a case-by-case basis. Analytical code is available upon request.

Conflicts of Interest

JL had received honoraria and served as a consultant or advisory board member from Otsuka, Janssen, Lundbeck, Sumitomo Pharmaceuticals, Boehringer Ingelheim, and ThoughtFull World Pte. Ltd. The other authors declare no conflicts of interest.

Multimedia Appendix 1
Transcription keys, interrater reliability, regression summaries, and model weights.
[DOCX File, 195 KB - formative_v9i1e75960_app1.docx]

References

1. Diagnostic and Statistical Manual of Mental Disorders: DSM-IV: American Psychiatric Association; 1994. [doi: 10.1176/ajp.152.8.1228]
2. Charlson FJ, Ferrari AJ, Santomauro DF, et al. Global epidemiology and burden of schizophrenia: findings from the global burden of disease study 2016. Schizophr Bull 2018 Oct 17;44(6):1195-1203. [doi: 10.1093/schbul/sby058] [Medline: 29762765]

3. Marder SR, Zito MF. Will I need to take these medications for the rest of my life? World Psychiatry 2018 Jun;17(2):165-166. [doi: 10.1002/wps.20519] [Medline: 29856554]
4. Fusar-Poli P, McGorry PD, Kane JM. Improving outcomes of first-episode psychosis: an overview. World Psychiatry 2017 Oct;16(3):251-265. [doi: 10.1002/wps.20446] [Medline: 28941089]
5. Huber G, Gross G. The concept of basic symptoms in schizophrenic and schizoaffective psychoses. Recenti Prog Med 1989 Dec;80(12):646-652. [Medline: 2697899]
6. Yung AR, McGorry PD. The prodromal phase of first-episode psychosis: past and current conceptualizations. Schizophr Bull 1996;22(2):353-370. [doi: 10.1093/schbul/22.2.353] [Medline: 8782291]
7. Miller TJ, McGlashan TH, Rosen JL, et al. Prodromal assessment with the structured interview for prodromal syndromes and the scale of prodromal symptoms: predictive validity, interrater reliability, and training to reliability. Schizophr Bull 2003;29(4):703-715. [doi: 10.1093/oxfordjournals.schbul.a007040] [Medline: 14989408]
8. Yung AR, Yuen HP, McGorry PD, et al. Mapping the onset of psychosis: the comprehensive assessment of at-risk mental states. Aust N Z J Psychiatry 2005;39(11-12):964-971. [doi: 10.1080/j.1440-1614.2005.01714.x] [Medline: 16343296]
9. Oliver D, Arribas M, Radua J, et al. Prognostic accuracy and clinical utility of psychometric instruments for individuals at clinical high-risk of psychosis: a systematic review and meta-analysis. Mol Psychiatry 2022 Sep;27(9):3670-3678. [doi: 10.1038/s41380-022-01611-w] [Medline: 35665763]
10. Bedi G, Carrillo F, Cecchi GA, et al. Automated analysis of free speech predicts psychosis onset in high-risk youths. NPJ Schizophr 2015;1(1):15030. [doi: 10.1038/npjschz.2015.30] [Medline: 27336038]
11. Cannon TD, Yu C, Addington J, et al. An individualized risk calculator for research in prodromal psychosis. Am J Psychiatry 2016 Oct 1;173(10):980-988. [doi: 10.1176/appi.ajp.2016.15070890] [Medline: 27363508]
12. Corcoran CM, Carrillo F, Fernández-Slezak D, et al. Prediction of psychosis across protocols and risk cohorts using automated language analysis. World Psychiatry 2018 Feb;17(1):67-75. [doi: 10.1002/wps.20491] [Medline: 29352548]
13. Fernandes BS, Karmakar C, Tamouza R, et al. Precision psychiatry with immunological and cognitive biomarkers: a multi-domain prediction for the diagnosis of bipolar disorder or schizophrenia using machine learning. Transl Psychiatry 2020 May 24;10(1):162. [doi: 10.1038/s41398-020-0836-4] [Medline: 32448868]
14. Koutsouleris N, Worthington M, Dwyer DB, et al. Toward generalizable and transdiagnostic tools for psychosis prediction: an independent validation and improvement of the NAPLS-2 risk calculator in the multisite PRONIA cohort. Biol Psychiatry 2021 Nov 1;90(9):632-642. [doi: 10.1016/j.biopsych.2021.06.023] [Medline: 34482951]
15. Mongan D, Föcking M, Healy C, et al. Development of proteomic prediction models for transition to psychotic disorder in the clinical high-risk state and psychotic experiences in adolescence. JAMA Psychiatry 2021 Jan 1;78(1):77-90. [doi: 10.1001/jamapsychiatry.2020.2459] [Medline: 32857162]
16. Rezaii N, Walker E, Wolff P. A machine learning approach to predicting psychosis using semantic density and latent content analysis. NPJ Schizophr 2019 Jun 13;5(1):9. [doi: 10.1038/s41537-019-0077-9] [Medline: 31197184]
17. Rodriguez-Ferrera S, McCarthy RA, McKenna PJ. Language in schizophrenia and its relationship to formal thought disorder. Psychol Med 2001 Feb;31(2):197-205. [doi: 10.1017/s003329170100321x] [Medline: 11232908]
18. Trzepacz PT, Baker RW. The Psychiatric Mental Status Examination: Oxford University Press; 1993. URL: https://catalog.nlm.nih.gov/discovery/fulldisplay/alma997224923406676/01NLM_INST:01NLM_INST [accessed 2025-12-06]
19. Covington MA, He C, Brown C, et al. Schizophrenia and the structure of language: the linguist's view. Schizophr Res 2005 Sep 1;77(1):85-98. [doi: 10.1016/j.schres.2005.01.016] [Medline: 16005388]
20. DeLisi LE. Speech disorder in schizophrenia: review of the literature and exploration of its relation to the uniquely human capacity for language. Schizophr Bull 2001;27(3):481-496. [doi: 10.1093/oxfordjournals.schbul.a006889] [Medline: 11596849]
21. Kuperberg GR. Language in schizophrenia part 1: an introduction. Lang Linguist Compass 2010 Aug;4(8):576-589. [doi: 10.1111/j.1749-818X.2010.00216.x] [Medline: 20936080]
22. Compton MT, Lunden A, Cleary SD, et al. The aprosody of schizophrenia: computationally derived acoustic phonetic underpinnings of monotone speech. Schizophr Res 2018 Jul;197:392-399. [doi: 10.1016/j.schres.2018.01.007] [Medline: 29449060]
23. Roche E, Segurado R, Renwick L, et al. Language disturbance and functioning in first episode psychosis. Psychiatry Res 2016 Jan 30;235:29-37. [doi: 10.1016/j.psychres.2015.12.008] [Medline: 26699880]
24. Bearden CE, Wu KN, Caplan R, Cannon TD. Thought disorder and communication deviance as predictors of outcome in youth at clinical high risk for psychosis. J Am Acad Child Adolesc Psychiatry 2011 Jul;50(7):669-680. [doi: 10.1016/j.jaac.2011.03.021] [Medline: 21703494]
25. Millman ZB, Goss J, Schiffman J, Mejias J, Gupta T, Mittal VA. Mismatch and lexical retrieval gestures are associated with visual information processing, verbal production, and symptomatology in youth at high risk for psychosis. Schizophr Res 2014 Sep;158(1-3):64-68. [doi: 10.1016/j.schres.2014.06.007] [Medline: 25000911]
26. Fusar-Poli P, Deste G, Smieskova R, et al. Cognitive functioning in prodromal psychosis: a meta-analysis. Arch Gen Psychiatry 2012 Jun;69(6):562-571. [doi: 10.1001/archgenpsychiatry.2011.1592] [Medline: 22664547]

27. Pawełczyk A, Łojek E, Żurner N, Kotlicka-Antczak M, Pawełczyk T. Higher order language impairments can predict the transition of ultrahigh risk state to psychosis-an empirical study. Early Interv Psychiatry 2021 Apr;15(2):314-327. [doi: 10.1111/eip.12943] [Medline: 32052573]
28. Spencer TJ, Thompson B, Oliver D, et al. Lower speech connectedness linked to incidence of psychosis in people at clinical high risk. Schizophr Res 2021 Feb;228:493-501. [doi: 10.1016/j.schres.2020.09.002] [Medline: 32951966]
29. Dalal TC, Liang L, Silva AM, Mackinley M, Voppel A, Palaniyappan L. Speech based natural language profile before, during and after the onset of psychosis: a cluster analysis. Acta Psychiatr Scand 2025 Mar;151(3):332-347. [doi: 10.1111/acps.13685] [Medline: 38600593]
30. Agurto C, Norel R, Wen B, et al. Are language features associated with psychosis risk universal? A study in Mandarin-speaking youths at clinical high risk for psychosis. World Psychiatry 2023 Feb;22(1):157-158. [doi: 10.1002/wps.21045] [Medline: 36640384]
31. Zhang H, He R, Palominos C, Hsu N, Cheung H, Hinzen W. The structure of meaning in schizophrenia: a study of spontaneous speech in Chinese. Psychiatry Res 2025 Feb;344:116347. [doi: 10.1016/j.psychres.2024.116347] [Medline: 39756103]
32. Natsuyama T, Chibaatar E, Shibata Y, et al. Associations of vocal features, psychiatric symptoms, and cognitive functions in schizophrenia. Neuropsychiatr Dis Treat 2025;21:943-954. [doi: 10.2147/NDT.S514927] [Medline: 40291596]
33. Mota NB, Ribeiro M, Malcorra B, et al. Attenuated symptoms are associated with connectedness and emotional expression in narratives based on emotional pictures in a Brazilian clinical high-risk cohort. Psychiatry Res 2025 Jun;348:116469. [doi: 10.1016/j.psychres.2025.116469] [Medline: 40174407]
34. Argolo F, Ramos WHDP, Mota NB, et al. Natural language processing in at-risk mental states: enhancing the assessment of thought disorders and psychotic traits with semantic dynamics and graph theory. Braz J Psychiatry 2024;46:e20233419. [doi: 10.47626/1516-4446-2023-3419] [Medline: 39074334]
35. Caplan R, Guthrie D, Fish B, Tanguay PE, David-Lando G. The kiddie formal thought disorder rating scale: clinical assessment, reliability, and validity. J Am Acad Child Adolesc Psychiatry 1989 May;28(3):408-416. [doi: 10.1097/00004583-198905000-00018] [Medline: 2738008]
36. Kuhn K, Kersken V, Reuter B, Egger N, Zimmermann G. Measuring the accuracy of automatic speech recognition solutions. ACM Trans Access Comput 2023 Dec 31;16(4):1-23. [doi: 10.1145/3636513]
37. Mujtaba D, Mahapatra N, Arney M, et al. Lost in transcription: identifying and quantifying the accuracy biases of automatic speech recognition systems against disfluent speech. Presented at: Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics; Jun 16-21, 2024. [doi: 10.18653/v1/2024.naacl-long.269]
38. Meehan K, McDermott F, Petropoulos N. Evaluating automatic transcription models utilising cloud platforms. 2024 Presented at: 2024 5th International Conference on Data Analytics for Business and Industry (ICDABI); Oct 23-24, 2024; Zallaq, Bahrain p. 91-96. [doi: 10.1109/ICDABI63787.2024.10800465]
39. Wollin-Giering S, Hoffmann M, Höfting J, Ventzke C. Automatic transcription of English and German qualitative interviews. Forum Qual Soc Res 2024;25(1). [doi: 10.17169/fqs-25.1.4129]
40. Alemu Y, Chen H, Duan C, Caulley D, Arriaga RI, Sezgin E. Detecting clinically relevant emotional distress and functional impairment in children and adolescents: protocol for an automated speech analysis algorithm development study. JMIR Res Protoc 2023 Jun 23;12:e46970. [doi: 10.2196/46970] [Medline: 37351936]
41. Min S, Shin D, Rhee SJ, et al. Acoustic analysis of speech for screening for suicide risk: machine learning classifiers for between- and within-person evaluation of suicidality. J Med Internet Res 2023 Mar 23;25:e45456. [doi: 10.2196/45456] [Medline: 36951913]
42. Riad R, Denais M, de Gennes M, et al. Automated speech analysis for risk detection of depression, anxiety, insomnia, and fatigue: algorithm development and validation study. J Med Internet Res 2024 Oct 31;26:e58572. [doi: 10.2196/58572] [Medline: 39324329]
43. Gibson CM, Penn DL, Prinstein MJ, Perkins DO, Belger A. Social skill and social cognition in adolescents at genetic risk for psychosis. Schizophr Res 2010 Sep;122(1-3):179-184. [doi: 10.1016/j.schres.2010.04.018] [Medline: 20570111]
44. Glenthøj LB, Kristensen TD, Gibson CM, Jepsen JRM, Nordentoft M. Assessing social skills in individuals at ultra-high risk for psychosis: validation of the High Risk Social Challenge Task (HiSoC). Schizophr Res 2020 Jan;215:365-370. [doi: 10.1016/j.schres.2019.08.025] [Medline: 31477371]
45. Lim K, Rapisarda A, Keefe RSE, Lee J. Social skills, negative symptoms and real-world functioning in individuals at ultra-high risk of psychosis. Asian J Psychiatr 2022 Mar;69:102996. [doi: 10.1016/j.ajp.2021.102996] [Medline: 35026654]
46. Addington J, Liu L, Perkins DO, Carrion RE, Keefe RSE, Woods SW. The role of cognition and social functioning as predictors in the transition to psychosis for youth with attenuated psychotic symptoms. Schizophr Bull 2017 Jan;43(1):57-63. [doi: 10.1093/schbul/sbw152] [Medline: 27798225]
47. Addington J, Penn D, Woods SW, Addington D, Perkins DO. Social functioning in individuals at clinical high risk for psychosis. Schizophr Res 2008 Feb;99(1-3):119-124. [doi: 10.1016/j.schres.2007.10.001] [Medline: 18023329]
48. Fusar-Poli P, Byrne M, Valmaggia L, et al. Social dysfunction predicts two years clinical outcome in people at ultra high risk for psychosis. J Psychiatr Res 2010 Apr;44(5):294-301. [doi: 10.1016/j.jpsychires.2009.08.016] [Medline: 19836755]

49. Lee J, Rekhi G, Mitter N, et al. The Longitudinal Youth at Risk Study (LYRIKS)—an Asian UHR perspective. Schizophr Res 2013 Dec;151(1-3):279-283. [doi: 10.1016/j.schres.2013.09.025] [Medline: 24139196]
50. Spiteri-Staines AE, Yung AR, Lin A, et al. Non-psychotic outcomes in young people at ultra-high risk of developing a psychotic disorder: a long-term follow-up study. Schizophr Bull 2024 Nov 8;50(6):1279-1286. [doi: 10.1093/schbul/sbae005] [Medline: 38366898]
51. Mitter N, Nah GQR, Bong YL, Lee J, Chong SA. Longitudinal Youth-at-Risk Study (LYRIKS): outreach strategies based on a community-engaged framework. Early Interv Psychiatry 2014 Aug;8(3):298-303. [doi: 10.1111/eip.12049] [Medline: 23682863]
52. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr Bull 1987;13(2):261-276. [doi: 10.1093/schbul/13.2.261] [Medline: 3616518]
53. Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: psychometric properties. J Consult Clin Psychol 1988;56(6):893-897. [doi: 10.1037/0022-006X.56.6.893]
54. Keefe RSE, Harvey PD, Goldberg TE, et al. Norms and standardization of the brief assessment of cognition in schizophrenia (BACS). Schizophr Res 2008 Jul;102(1-3):108-115. [doi: 10.1016/j.schres.2008.03.024] [Medline: 18495435]
55. VLC media player. VideoLan. 2006. URL: https://www.videolan.org/vlc/index.html [accessed 2025-02-14]
56. Boersma P, Weenink D. Praat. 2024. URL: http://www.praat.org/ [accessed 2024-04-14]
57. Jacewicz E, Fox RA, O'Neill C, Salmons J. Articulation rate across dialect, age, and gender. Lang Var Change 2009 Jul 1;21(2):233-256. [doi: 10.1017/S0954394509990093] [Medline: 20161445]
58. Honnibal M, Montani I, Van Landeghem S, Boyd A. spaCy: industrial-strength natural language processing in Python. GitHub. 2020. URL: https://github.com/explosion/spaCy/blob/master/CITATION.cff [accessed 2025-12-18]
59. Pauselli L, Halpern B, Cleary SD, Ku BS, Covington MA, Compton MT. Computational linguistic analysis applied to a semantic fluency task to measure derailment and tangentiality in schizophrenia. Psychiatry Res 2018 May;263:74-79. [doi: 10.1016/j.psychres.2018.02.037] [Medline: 29502041]
60. Voppel AE, de Boer JN, Brederoo SG, Schnack HG, Sommer I. Quantified language connectedness in schizophrenia-spectrum disorders. Psychiatry Res 2021 Oct;304:114130. [doi: 10.1016/j.psychres.2021.114130] [Medline: 34332431]
61. Glasgow K, Roos M, Haufler A, Chevillet M, Wolmetz M. Evaluating semantic models with word-sentence relatedness. arXiv. Preprint posted online on Mar 23, 2016. [doi: 10.48550/arXiv.1603.07253]
62. Villegas MP, Garciarena Ucelay MJ, Fernández JP, Álvarez Carmona MA, Errecalde ML, Cagnina L. Vector-based word representations for sentiment analysis: a comparative study. Presented at: XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016) [22nd Argentine Congress of Computer Science]; Oct 3-7, 2016 URL: http://sedici.unlp.edu.ar/handle/10915/56763 [accessed 2016-11-16]
63. Gelfer MP, Mikos VA. The relative contributions of speaking fundamental frequency and formant frequencies to gender identification based on isolated vowels. J Voice 2005 Dec;19(4):544-554. [doi: 10.1016/j.jvoice.2004.10.006] [Medline: 16301101]
64. Murray IR, Arnott JL. Toward the simulation of emotion in synthetic speech: a review of the literature on human vocal emotion. J Acoust Soc Am 1993 Feb;93(2):1097-1108. [doi: 10.1121/1.405558] [Medline: 8445120]
65. Laukka P, Linnman C, Åhs F, et al. In a nervous voice: acoustic analysis and perception of anxiety in social phobics' speech. J Nonverbal Behav 2008 Dec;32(4):195-214. [doi: 10.1007/s10919-008-0055-9]
66. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B 1995 Jan 1;57(1):289-300. [doi: 10.1111/j.2517-6161.1995.tb02031.x]
67. Iyortsuun NK, Kim SH, Jhon M, Yang HJ, Pant S. A review of machine learning and deep learning approaches on mental health diagnosis. Healthcare (Basel) 2023 Jan 17;11(3):285. [doi: 10.3390/healthcare11030285] [Medline: 36766860]
68. Hastie T, Tibshirani R, Friedman JH. The Elements of Statistical Learning: Data Mining, Inference, and Prediction: Springer; 2009. [doi: 10.1007/978-0-387-84858-7]
69. Musa AB. Comparative study on classification performance between support vector machine and logistic regression. Int J Mach Learn Cyber 2013 Feb;4(1):13-24. [doi: 10.1007/s13042-012-0068-x]
70. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine learning in Python. J Mach Learn Res 2012 Jan 2;12:2825-2830 [FREE Full text]
71. Platt J. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In: Advances in Large-Margin Classifiers: MIT Press; 1999:61-73. [doi: 10.7551/mitpress/1113.003.0008]
72. Chicco D, Tötsch N, Jurman G. The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. BioData Min 2021 Feb;14(1):13. [doi: 10.1186/s13040-021-00244-z] [Medline: 33541410]
73. Parola A, Simonsen A, Bliksted V, Fusaroli R. Voice patterns in schizophrenia: a systematic review and Bayesian meta-analysis. Schizophr Res 2020 Feb;216:24-40. [doi: 10.1016/j.schres.2019.11.031] [Medline: 31839552]
74. Corcoran CM, Mittal VA, Bearden CE, et al. Language as a biomarker for psychosis: a natural language processing approach. Schizophr Res 2020 Dec;226:158-166. [doi: 10.1016/j.schres.2020.04.032] [Medline: 32499162]
75. van Rooijen G, Isvoranu AM, Meijer CJ, et al. A symptom network structure of the psychosis spectrum. Schizophr Res 2017 Nov;189:75-83. [doi: 10.1016/j.schres.2017.02.018] [Medline: 28237606]

76. Piskulic D, Addington J, Cadenhead KS, et al. Negative symptoms in individuals at clinical high risk of psychosis. Psychiatry Res 2012 Apr 30;196(2-3):220-224. [doi: 10.1016/j.psychres.2012.02.018] [Medline: 22445704]
77. Shin YS, Kim SY, Lee TY, et al. Longitudinal change in neurocognition and its relation to symptomatic and functional changes over 2 years in individuals at clinical high-risk for psychosis. Schizophr Res 2016 Jul;174(1-3):50-57. [doi: 10.1016/j.schres.2016.03.024] [Medline: 27068568]
78. Hitczenko K, Mittal VA, Goldrick M. Understanding language abnormalities and associated clinical markers in psychosis: the promise of computational methods. Schizophr Bull 2021 Mar 16;47(2):344-362. [doi: 10.1093/schbul/sbaa141] [Medline: 33205155]

Abbreviations

AR: articulation rate
ASR: automatic speech recognition
BA: balanced accuracy
BACS: Brief Assessment of Cognition in Schizophrenia
BAI: Beck Anxiety Inventory
DD: depression diagnosis
DF: dysfluency
GAI: generative artificial intelligence
HiSoC: High-Risk Social Challenge
LSA: latent semantic analysis
LYRIKS: Longitudinal Youth at Risk Study
MCC: Matthew correlation coefficient
NLP: natural language processing
PANSS: Positive and Negative Syndrome Scale
PSLE: Primary School Leaving Examination
SC: sequential coherence
SVM: support vector machine
TMT: token motor task
UHR: ultra-high risk
WPM: words per minute

Edited by J Sarvestan; submitted 14.04.25; peer-reviewed by A Loch, D Patel, E Stroulia, L Xu; revised version received 11.11.25; accepted 12.11.25; published 30.12.25.

Please cite as:
Tan SMX, Lieu MY, Kai J, Yang Z, KK L, Lwin MO, Lee J, Goh WWB
Predicting Ultra-High Risk Outcomes Using Linguistic and Acoustic Measures From High-Risk Social Challenge Recordings: mHealth Longitudinal Cohort Exploratory Study
JMIR Form Res 2025;9:e75960
URL: https://formative.jmir.org/2025/1/e75960
doi:10.2196/75960

Original Paper

Prescribing Experiences, Potentials, and Challenges of Digital Health Applications in the Field of Hormones and Metabolism: Cross-Sectional Survey Study of Health Care Providers in Germany

Melanie Mäder[1,2]; Dirk Müller-Wieland[3], Prof Dr; Tobias Wiesner[4], Dr; Tonio Schoenfelder[2], Dr; Carsta Militzer-Horstmann[1,2], Dr; Ria Heinrich[2]; Mareike Geisler[2], Dr; Dennis Häckl[1,2], Jr Prof Dr

[1]Chair for Health Economics and Management, Faculty of Economics and Management Science, Leipzig University, Leipzig, Germany

[2]Scientific Institute for Health Economics and Health System Research, Leipzig, Germany

[3]Department of Medicine I, Universitätsklinikum Aachen, Aachen, Germany

[4]Medical Care Center for Metabolic Medicine, Leipzig, Germany

Corresponding Author:
Melanie Mäder
Chair for Health Economics and Management
Faculty of Economics and Management Science
Leipzig University
Grimmaische Straße 12
Leipzig, 04109
Germany
Phone: 49 17682289222
Email: melanie.maeder@wig2.de

Abstract

Background: In 2020, the global prevalence of overweight and obesity was approximately 42%. One of the most common associated conditions is type 2 diabetes mellitus, which had a global prevalence of around 10.5% in 2021. Digital health applications (DiHA), which can be prescribed as certified medical devices in Germany, have been shown to effectively support disease management in patients with overweight and diabetes mellitus. However, little is known about DiHA-prescribing behavior of health care providers (HCPs) specializing in hormones and metabolism or about potential barriers to prescribing these applications.

Objective: This study aimed to assess HCPs' experience with and willingness to prescribe DiHA in the field of hormones and metabolism. In addition, it sought to examine the patient-relevant health care effects that HCPs perceive as potentially achievable or have already observed with DiHA use, as well as the barriers they perceive to prescribing these applications.

Methods: An online questionnaire was developed based on preliminary studies and a literature review consisting of 86 items covering 6 key areas: experience and willingness to prescribe, health care effects, barriers, scientific evidence, digital affinity, and sociodemographics. The anonymous survey was distributed via the German Diabetes Association to 6035 HCPs in Germany between August 2 and October 9, 2024. Descriptive analyses, as well as correlation and regression analyses, were conducted.

Results: A total of 350 HCPs participated in the survey (response rate=5.8%). Although the low response rate may limit generalizability, the findings provide insights into prescribing behavior within this specialty. More than half (187/350, 53.4%) had never prescribed any of the 54 DiHA available at the time of the survey, with 47.6% (89/187) citing a lack of experience as the primary reason. Among those who had prescribed a DiHA (163/350, 46.6%), the majority (139/163, 85.3%) had prescribed 1 of the 8 DiHA available for obesity or diabetes mellitus. Looking ahead, 42.9% (149/348) of all surveyed HCPs stated that they were either very unlikely (83/348, 23.9%) or somewhat unlikely (66/348, 19%) to prescribe these DiHA in the next 12 months. The greatest perceived benefits of DiHA were improvements in self-management, health literacy, and adherence. The main barriers to prescribing DiHA in the field of hormones and metabolism included inadequate reimbursement for ancillary medical services, poor compatibility with existing practice software, and a lack of digital affinity or motivation among patients.

Conclusions: DiHA have not yet been fully integrated into standard health care. To improve prescribing, we recommend integrating DiHA into medical guidelines, ensuring proper reimbursement, and involving HCPs in the pricing and health-economic

evaluation of DiHA. The recommendations outlined should be considered to maximize DiHA's potential and improve HCPs' acceptance, providing valuable insights for health policy to enhance the integration, reimbursement, and use of DiHA.

(JMIR Form Res 2025;9:e77792) doi:10.2196/77792

KEYWORDS

chronic disease management; diabetes mellitus type 2; digital health applications; health care providers; medical decision-making; obesity; overweight

Introduction

Digital interventions provide substantial opportunities for enhancing the management of chronic conditions [1], particularly by strengthening patient empowerment, supporting self-management [2], and improving treatment adherence [3]. They are especially relevant in the context of chronic diseases such as overweight, obesity, and type 2 diabetes mellitus (T2DM), which are associated with a significant clinical and health-economic burden [4,5].

In 2020, approximately 42% of the global population was classified as overweight (n=1.39 billion) or obese (n=0.81 billion), and this proportion is projected to rise to 54% by 2035, representing an estimated 1.77 billion individuals with overweight and 1.53 billion individuals with obesity [5]. T2DM is among the most prevalent secondary conditions associated with obesity, with a global prevalence of approximately 10.5% in 2021, projected to rise to 12.2% by 2045 [4]. Both obesity and diabetes mellitus (DM) are associated with significant impairments in quality of life (QoL) [6-8].

Digital interventions have the potential to enhance clinical outcomes and improve QoL in individuals with obesity or DM [9]. In DM applications, this is primarily reflected in clinically significant reductions in glycated hemoglobin (HbA_{1c}) levels [10,11], while in obesity applications, it is demonstrated by clinically significant reductions in body weight [12,13]. Digital health applications (DiHA) were introduced into the German health care system in 2019 as components of standard care covered by statutory health insurance (SHI) funds. Applications are eligible for DiHA status if they successfully complete the evaluation procedure conducted by the Federal Institute for Drugs and Medical Devices (Bundesinstitut für Arzneimittel und Medizinprodukte [BfArM]) and demonstrate a measurable benefit compared with standard therapy. As of December 7, 2025, the DiHA directory comprises 57 registered applications. Following a confirmed diagnosis (*ICD-10* [*International Statistical Classification of Diseases, Tenth Revision*] code), these applications can be prescribed by physicians or psychotherapists, underscoring the essential role of these professional groups in the deployment of DiHA and the advancement of a digitally enabled health care system.

Digital interventions like DiHA, and their acceptance and use by health care providers (HCPs), have been extensively investigated through quantitative surveys in Germany which indicate that HCP's willingness to prescribe them remains low. A nationwide survey in 2020 found that 37% of general practitioners (GPs; N=51) reported a rather high or high willingness to prescribe DiHA [14]. Another study conducted between December 2020 and January 2021 (N=1308) revealed that only 30.3% of HCPs planned to continue prescribing DiHA in the future, despite acknowledging their benefits [15]. A 2022 survey among GPs (N=3829) in 4 German states found an even lower willingness to prescribe at just 13% [16]. In contrast, a nationwide study of diabetes specialists (N=538) conducted between September 2021 and April 2022 showed that 51% perceived health apps as potentially beneficial for T2DM, though further research was needed to evaluate their strengths and weaknesses [17]. A more recent survey in 2023 reported that 38.5% of GPs (N=97) in Giessen were very likely to prescribe DiHA in the next 12 months, highlighting the ongoing need for education and information to improve acceptance [18]. The studies mentioned indicate significant room for improvement in prescribing willingness, but all these studies either focus on DiHA use and acceptance by GPs in a broad context, without addressing specific indication areas, or they examine digital interventions in general, without specifically focusing on DiHA. Focusing the survey on a specific indication area—hormones and metabolism in this case—is crucial to ensure the validity and relevance of the results. DiHA vary widely in terms of application profiles, levels of evidence, acceptance, and clinical relevance across different medical specialties. A generalized perspective across all indications would obscure these differences and prevent meaningful conclusions from being drawn regarding the use, evaluation, or implementation of DiHA within a specific health care context. Therefore, to better understand indication-specific prescribing experiences, potential benefits, and challenges associated with DiHA, it is necessary to analyze DiHA adoption within individual medical specialties. So, the planned research project focuses on prescribing experiences, potential benefits, and challenges associated with DiHA within the field of hormones and metabolism—hereinafter referred to as "hormones and metabolism digital health applications" (HM-DiHA)—and aims to address the following research questions:

Our primary questions are as follows:

1. What is the experience and willingness of HCPs to prescribe HM-DiHA?
2. What patient-relevant health care effects can potentially be achieved and have already been achieved using HM-DiHA?
3. What barriers do HCPs perceive regarding the prescription of HM-DiHA?
4. What actionable recommendations can be derived from the findings to enhance HCPs' acceptance of HM-DiHA?

Our secondary question is as follows:

Is there a correlation between HCPs' (1) previous experience with DiHA prescription, (2) prescription frequency, and (3)

intention to prescribe HM-DiHA and the following characteristics: gender, age, medical specialization, activity within the framework of SHI care, additional professional qualification (eg, diabetologist, adiposiologist, etc) professional experience, location of practice (federal state and practice environment), working model, number of patients treated per quarter, and digital affinity?

Methods

Study Design

To address the predefined research questions, the authors used a cross-sectional study design to gather comprehensive opinions and experiences from HCPs in the field of hormones and metabolism; the Ethics Advisory Board of Leipzig University reviewed the application for the research project using a simplified procedure and concluded that there were no ethical objections to its implementation. In a rapidly evolving digital health care landscape, this design offers valuable insights into the current levels of acceptance and implementation of HM-DiHA in clinical practice. With the support of the German Diabetes Association (GDA), a nationwide quantitative online survey was conducted in Germany between August and October 2024. The GDA is a medical-scientific professional society with approximately 9300 members, including physicians in clinics and practices, scientists, psychologists, pharmacists, and other specialists in the field of diabetology [19]. Recruiting participants through the GDA proved to be effective, as its members possess a high level of expertise in managing chronic conditions such as obesity or DM—an area in which DiHA are particularly relevant. Targeting this specialized group ensured both the thematic relevance of the survey and a practice-oriented assessment of DiHA usage and acceptance.

Survey Instrument

The questionnaire was designed based on preliminary studies by the authors and an extensive literature search on the individual key topics: (1) experiences and willingness to prescribe (14 items) [14-16], (2) health care effects (potential and actual experience; 17 items each) [20], (3) barriers (11 items) [21], (4) study evidence (7 items) [22], (5) digital affinity (4 items), and (6) sociodemographics (16 items). The questionnaire thus comprised a total of 86 items. The number of items answered varied depending on prior responses (eg, blocks were skipped if no DiHA had been prescribed).

Ordinal scales were predominantly used to operationalize individual theme blocks (especially 2-4) to allow the HCPs to answer intuitively while ensuring a sufficient level of information and the highest data quality possible. These ordinal scales were unipolar rating scales, designed to represent the graduated characteristics of individual items. All scale markers were expressed through verbal labels using linguistically clear and widely recognized formulations [23] to facilitate quick responses. A 5-point rating scale was deliberately chosen, as scales with 5 to 7 levels have proven effective and offer optimal psychometric properties [23]. The use of an odd-numbered scale ensured the inclusion of a neutral middle category to capture potential uncertainties in assessments. Additionally, participants had the option to select “I don’t know,” which explicitly indicated difficulties in making an assessment. In some cases, nominal and interval scales were also applied. Alongside standardized questions, open-ended questions were included (eg, regarding problems and barriers to prescribing). The full translated questionnaire is available in Multimedia Appendix 1. The questionnaire was refined through multistage pretesting, including expert reviews and a pilot with 5 HCPs under field conditions; however, no formal psychometric validation (eg, assessment of internal consistency or reliability testing) of the instrument was conducted.

Recruitment and Sample

Between August 2 and October 9, 2024, the GDA invited 6035 HCPs via email, sent through a mailing list, to participate in an anonymous online survey, with a reminder sent on September 26. The survey targeted all GDA-registered HCPs in the field of hormones and metabolism eligible to prescribe DiHA, representing a full census of this group.

Data Analysis

Data analysis was performed using SPSS software (version 23; IBM Corp). We presented data using descriptive statistical methods such as frequency, cross-tables and bar charts, as well as individual characteristic values. These included measures of central tendency, such as the mean and median, and measures of dispersion, such as SD, IQR, and range [23].

Bivariate correlation analysis was used to assess both the strength (weak vs strong) and, where applicable, the direction (positive vs negative) of correlations. Depending on the scale level of the variables under investigation, we applied different correlation coefficients (r):

- Nominal × Nominal or Nominal × Ordinal (Cramér V): $0.1 \le V \le 0.29$ small effect size; $0.3 \le V \le 0.49$ medium effect size; $V \ge 0.5$ large effect size [24]
- Nominal × Metric (eta coefficient, η; 0=no correlation, 1=perfect correlation): $0.100 \le \eta \le 0.242$ small effect size; $0.243 \le \eta \le 0.370$ medium effect size; $\eta \ge 0.371$ large effect size [24]
- Ordinal × Ordinal or Ordinal × Metric (Spearman rank correlation coefficient, ρ; –1=perfect negative correlation, 0=no correlation, +1=perfect positive correlation): $0.1 \le \rho \le 0.29$ small effect size; $0.3 \le \rho \le 0.49$ medium effect size; $\rho \ge 0.5$ large effect size [24]

The significance level (α) was set at 5% (ie, α=.05) [23].

Binary logistic regression analysis was used to examine the relationship between the dependent variable probability—assuming the value 1—and the independent variables. This method tested whether a relationship exists between the independent variables and a binary dependent variable. The dependent variable was defined as the general DiHA prescription (no=0, yes=1). Independent variables were selected based on the literature and were only included in the regression analysis if they met the following criteria [23]:

- For each group formed by categorical predictors, $n \ge 25$.
- Independent variables were not highly correlated with each other ($r<0.7$).

Since the regression model was based on well-founded theoretical considerations, we applied the "inclusion" method, meaning all variables within a block were entered into the model in a single step. All results are presented as odds ratios (ORs) with 95% CI.

Ethical Considerations

The Ethics Advisory Board of Leipzig University reviewed the application for the research project using a simplified procedure and concluded that there were no ethical objections to its implementation.

Results

The questionnaire was distributed to a total of 6035 HCPs, of whom 288 completely and 140 partially responded. After verifying and cleaning the data, 78 cases were excluded because no information was provided; the questionnaire was clicked through without being filled out. Ultimately, 350 HCPs participated in the survey, resulting in a response rate of 5.8%.

Sociodemographics

Of the respondents, 45.9% (133/290) were male, and 53.1% (154/290) were female. Around a quarter of participants were aged 36-45 years (72/290, 24.8%), 46-55 years (86/290, 29.7%), and 56-65 years (82/290, 28.3%). The majority (278/290, 95.9%) were specialists, primarily in internal medicine (196/290, 63.6%). In terms of workplace settings, 42.5% (99/290) worked in general practice, and 35.6% (83/290) in specialist care. Additionally, 82.4% (238/289) held an additional qualification, most commonly in diabetology (63.8%). The largest group worked in hospitals (84/290, 29.3%), followed by group practices (80/290, 27.6%; Multimedia Appendix 2).

Digital Affinity

HCPs rated their own digital affinity on a scale from 0 (not at all digitally affine) to 10 (very digitally affine), with an average rating of 6.80 (SD 2.03). Respondents who had already prescribed a DiHA or an HM-DiHA rated their own digital affinity with an average of 6.92 (SD 1.80) and 6.96 (SD 1.75), respectively, slightly higher compared to those who had not yet prescribed a DiHA or a DiHA from the indication area with 6.69 (SD 2.11) and 6.70 (SD 2.08). There was a correlation with a small effect size between general prescribing experience (η=0.057) and experience with prescribing HM-DiHA (η=0.051) and the respondents' self-assessed digital affinity. HCPs' own digital affinity did not correlate significantly with the intention to prescribe (ρ=0.156; P=.08).

Of the 292 respondents, 51.4% (150/292) had never used a health app as a patient, and this proportion was even higher for DiHA, with 86.6% (253/292) having never used one. A total of 34.6% (101/292) of HCPs had accessed manufacturer-provided test versions to evaluate a DiHA before prescribing it to patients. Among HCPs who prescribed a DiHA in general, more than half (75/137, 54.7%) had used manufacturer access, while 45.3% (62/137) had not; the correlation between general DiHA prescription and manufacturer access use was statistically significant (V=0.398; P<.001; Multimedia Appendix 3), with regression analysis further revealing that HCPs who had used manufacturer access were almost 12 times more likely to have prescribed a DiHA compared to those who had not (Table 1).

A statistically significant correlation with medium effect size was also found between an HM-DiHA prescription and manufacturer access use (V=0.330; P<.001), with 61.5% (72/117) of HCPs who prescribed a DiHA from this indication area having also used manufacturer access (Multimedia Appendix 3). There was a statistically significant correlation with medium effect size between the use of a DiHA as a patient (V=0.193; P=.05) and the use of manufacturer access (V=0.337; P<.001) with the intention to prescribe, with 30.9% (17/55) of HCPs who were more likely and 61.2% (41/67) who were likely to prescribe an HM-DiHA in the next 12 months having already used manufacturer access (Multimedia Appendix 4).

Table 1. Multivariate representation of the dependent variable "general digital health applications (DiHA) prescription," categorized by independent variables. Details on assumption checks for the binary logistic regression are provided in Multimedia Appendix 6.

Independent variables	Regression coefficient (B)	*P* value	Odds ratio (95% CI)
Gender (reference=male)			
Female	1.053	.01	2.867 (1.235-6.655)
Activity within the framework of statutory health insurance care (reference=general practitioner care)			
Specialist care	–1.454	.002	0.234 (0.093-0.587)
Additional title (reference=no)			
Yes	0.322	.58	1.38 (0.441-4.321)
Patients treated per quarter (reference=<500)			
500-750	1.061	.16	2.89 (0.662-12.616)
751-1000	1.548	.03	4.704 (1.148-19.274)
1001-1500	2.447	.002	11.554 (2.472-54.012)
1501-2000	1.153	.17	3.169 (0.616-16.301)
>2000	2.138	.008	8.48 (1.738-41.384)
Ever used a health app as a patient (reference=no)			
Yes	0.224	.63	1.251 (0.504-3.106)
Ever used a DiHA as a patient (reference=no)			
Yes	–0.206	.76	0.814 (0.219-3.030)
Ever used a DiHA manufacturer access (reference=no)			
Yes	2.459	<.001	11.689 (4.370-31.266)
Digital affinity (0=not at all digitally affine to 10=very digitally affine)	0.001	.99	1.001 (0.818-1.226)
Omnibus test of the model coefficients	N/A[a]	<.001	N/A
Cox and Snell R^2	N/A	.37	N/A
Nagelkerke R^2	N/A	.49	N/A

[a]N/A: not applicable.

Experience With Previous DiHA Prescription

General DiHA Prescription

At 53.4% (187/350), more than half of the respondents had not yet prescribed a DiHA to their patients, with almost half of the respondents (89/187, 47.6%) named a lack of experience as the reason (multiple selection was possible). About one-fifth cited barriers such as insufficient efficacy evidence, low patient digital literacy, lack of interest in prescribing DiHA among HCPs, integration issues, time constraints, additional time requirements, and lack of training opportunities (Multimedia Appendix 5).

As part of an open-ended question, HCPs had the opportunity to provide additional reasons for not prescribing DiHA. The most common reasons included treating children and adolescents (15/64) for whom DiHA are not yet approved, or working in an inpatient setting (13/64), where prescribing a DiHA is either not possible or deliberately avoided due to the inability to ensure follow-up care. A lack of digital affinity, as well as insufficient information and knowledge (8/64), were also named.

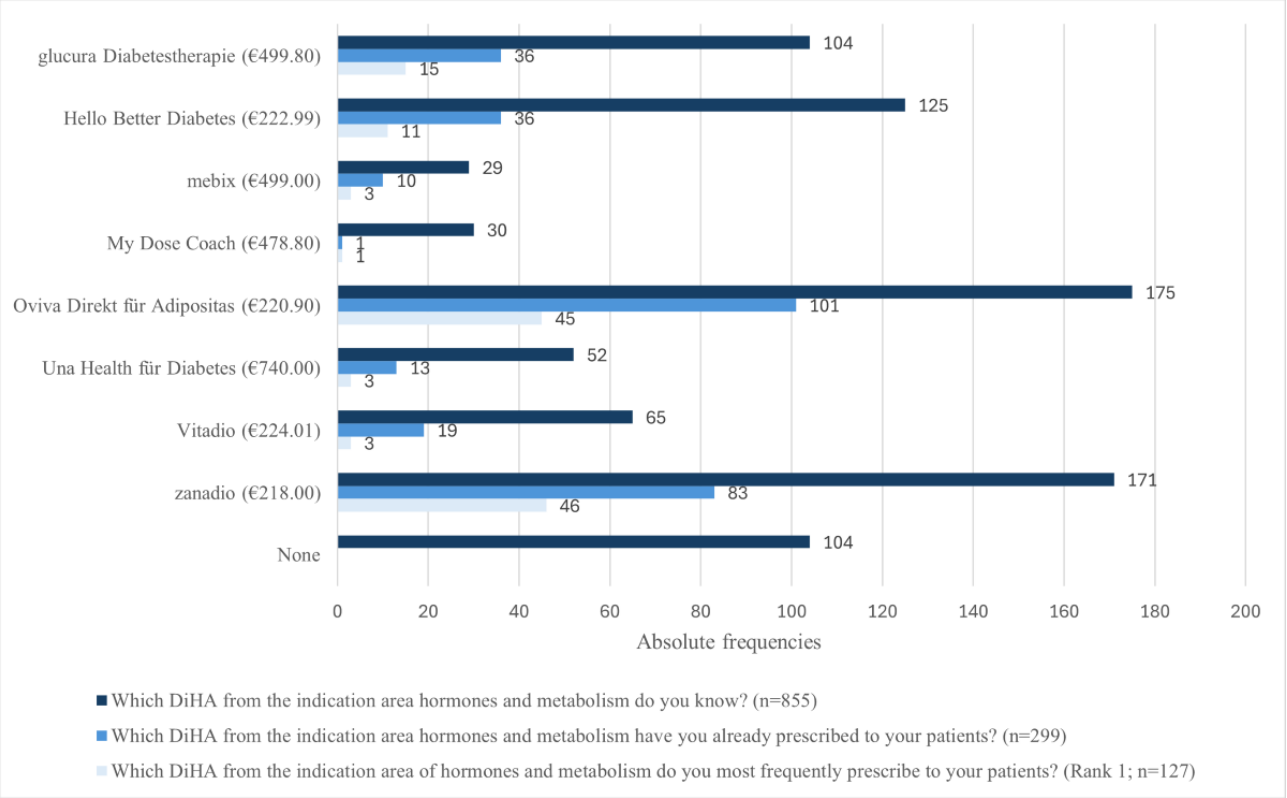
Regarding the general DiHA prescription, statistically significant correlations with small to medium effect sizes were observed for the activity within the framework of SHI (V=0.481; P<.001), the use of an additional professional title (V=0.142; P=.05), the size of the city or municipality (V=0.231; P=.009), the working model (V=0.437; P<.001), and the number of patients treated per quarter (V=0.477; P<.001; Multimedia Appendix 3). Women were nearly 3 times more likely than men to have already prescribed a DiHA (OR 2.87, 95% CI 1.235-6.655; P=.01). HCPs treating 1001 to 1500 patients per quarter had 11.6 times higher odds (OR 11.55, 95% CI 2.472-54.012; P=.002) of having prescribed a DiHA, while those treating more than 2000 patients per quarter had 8.5 times higher odds (OR 8.48, 95% CI 1.738-41.384; P=.008), compared to HCPs treating fewer than 500 patients per quarter (Table 1).

Indication-Specific DiHA Prescription

Of the HCPs who already prescribed a DiHA, the majority (139/163, 85.3%) had prescribed an HM-DiHA, while 14.7% (24/163) had not. Most HCPs were familiar with and frequently prescribed DiHA Oviva Direkt für Adipositas (n=175 familiar; n=101 prescribed) and zanadio (n=171 familiar; n=83 prescribed), while mebix (n=29 familiar; n=10 prescribed; the DiGA was removed from the DiGA directory on July 14, 2025, as no positive health care effect could be demonstrated) and My

Dose Coach (n=30 familiar; n=1 prescribed; the DiGA My Dose Coach was removed from the DiGA directory on January 10, 2026, at the manufacturer's request) were less known and infrequently prescribed. Overall, 104 HCPs were unaware of any DiHA in this area (Figure 1).

Figure 1. Experience with specific digital health applications (DiHA).

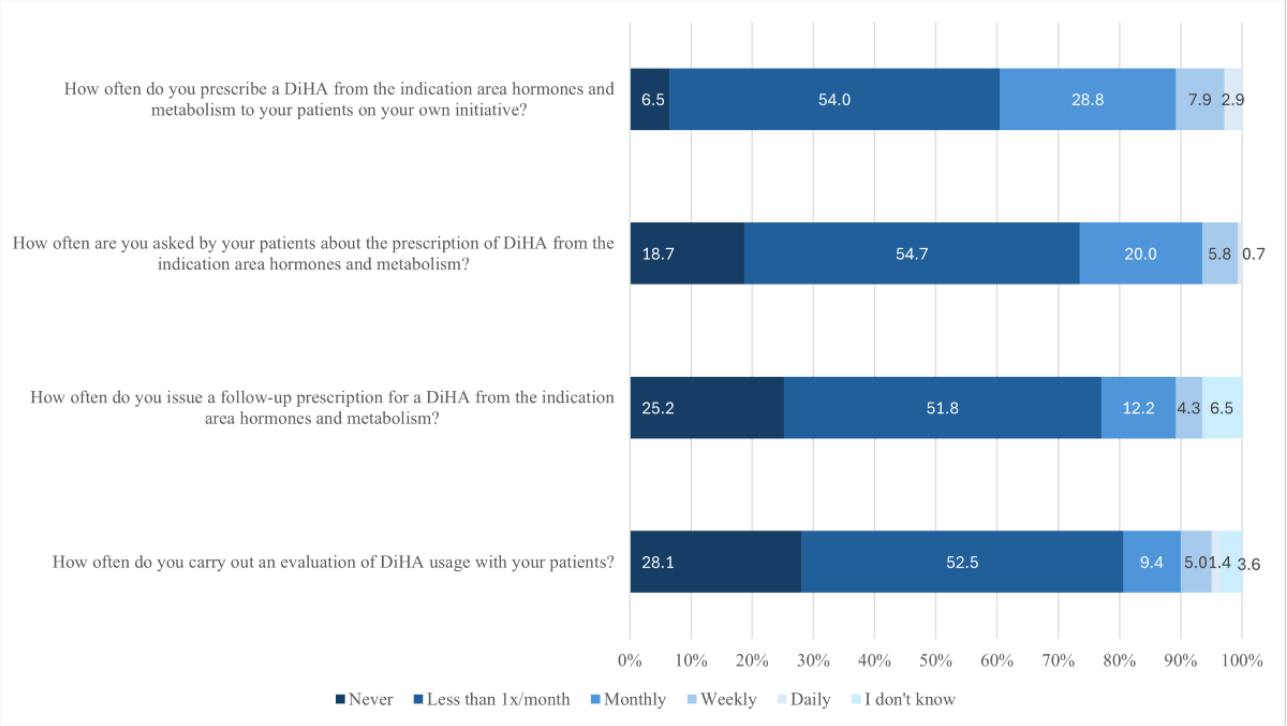


DiHA were most commonly prescribed for patients aged 36-45 years (n=54) and considered appropriate for all age groups, particularly 36-45 years (n=284) and 26-35 years (n=273), while least prescribed for those older than 65 years (n=4) and younger than 25 years (n=8).

Prescription Frequency

More than half of the respondents had never (9/139, 6.5%) or less than once a month (75/139, 54%) prescribed an HM-DiHA on their own initiative. Similarly, three-quarters of HCPs had never been asked by patients to prescribe an HM-DiHA (26/139, 18.7%) or were asked less than once a month (76/139, 54.7%) for such a prescription. Follow-up prescriptions were also issued infrequently, with 35/139 (25.2%) never prescribing a follow-up and 72/139 (51.8%) prescribing them less than once a month (Figure 2). No statistically significant correlation was found between the frequency of prescribing DiHA on the HCPs' own initiative and all the sociodemographic variables (Multimedia Appendix 7).

Figure 2. Experience with prescription frequency of digital health applications (DiHA).



Prescription Intention

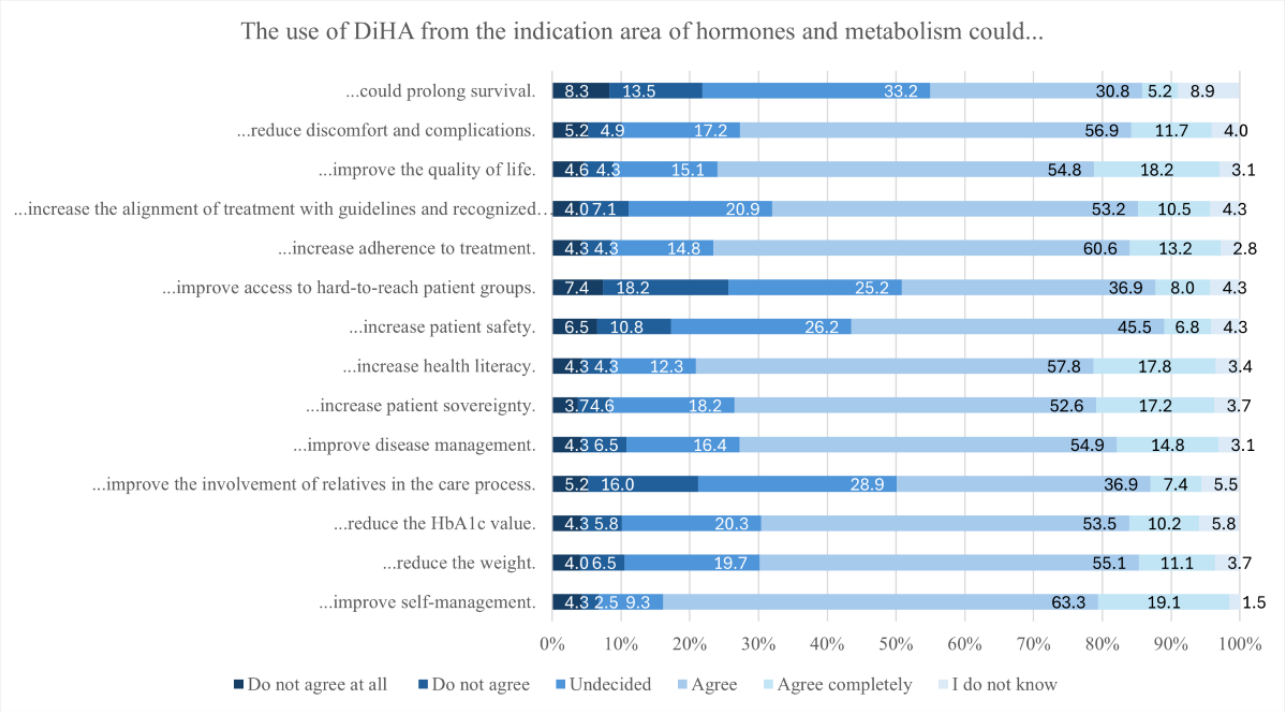
Overall, 42.9% (149/348) of respondents were either very unlikely (83/348, 23.9%) or somewhat unlikely (66/348, 19%) to prescribe HM-DiHA in the next 12 months, while 40.5% (141/348) were either very likely (80/348, 23%) or somewhat likely (61/348, 17.5%) to do so. Additionally, 12.6% (44/348) of HCPs were undecided.

We found a statistically significant correlation between the intention to prescribe a DiHA in the next 12 months and activity within the framework of SHI (V=0.266; P<.001), with the majority (38/56, 67.9%) of those who were very likely to prescribe a DiHA being GPs, while only 23.2% (13/56) of specialists expressed the same likelihood. A statistically significant correlation was also identified between the intention to prescribe and the working model (V=0.246; P<.001), with 37% (20/54) of those rather unlikely and 40.3% (27/67) of those very unlikely to prescribe a DiHA working in hospitals (Multimedia Appendix 4).

Health Care Effects

A total of 325 HCPs evaluated the potential positive health care effects that HM-DiHA could offer. The majority of respondents (267/325, 82.4%) indicated that these applications could enhance self-management (agree: 205/325, 63.3%; strongly agree: 62/325, 19.1%). Three-quarters of respondents believed that DiHA could improve health literacy, adherence, and QoL. The areas where the least potential for improvement was noted were in prolonging survival (agree: 100/325, 30.8%; strongly agree: 17/325, 5.2%), increasing the involvement of relatives in the treatment process (agree: 120/325, 36.9%; strongly agree: 24/325, 7.4%), and improving access to hard-to-reach patient groups (agree: 120/325, 36.9%; strongly agree: 26/325, 8%; Figure 3; Multimedia Appendix 8).

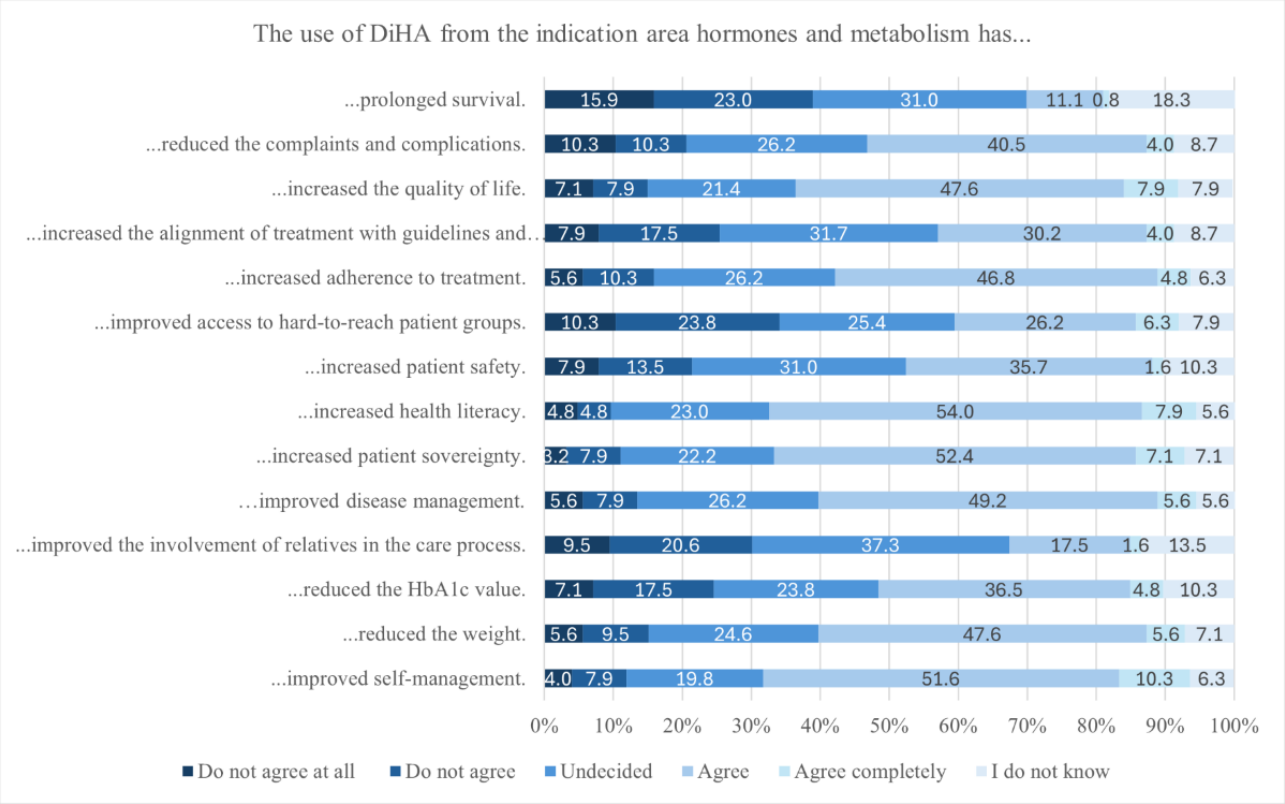
Figure 3. Potentially achievable positive health care effects of digital health applications (DiHA).



Apart from prolonging survival, improving access to hard-to-reach patient groups, and enhancing self-management, we found weak statistically significant correlations between individual potential positive health care effect assessment and experience with general DiHA prescriptions. Additionally, statistically significant correlations were observed between the assessment of potential positive health care effects—except for the involvement of relatives in the care process and improvement of self-management—and the frequency of prescriptions, as well as between the assessment of potential positive health care effects and the intention to prescribe (Multimedia Appendix 9).

A total of 126 HCPs evaluated the positive health care effects that HM-DiHA have already achieved for their patients. Most respondents (78/126, 61.9%) reported that these applications had improved their patients' self-management (agree: 65/126, 51.6%; strongly agree: 13/126, 10.3%) or health literacy (agree: 68/126, 54%; strongly agree: 10/126, 7.9%). More than half of the HCPs also observed improvements in their patients' QoL, adherence, patient autonomy, disease management, and weight. The least positive health care effects were noted in relation to prolonging survival (agree: 14/126, 11.1%; strongly agree: 1/126, 0.8%) and increasing the involvement of relatives in the care process (agree: 22/126, 17.5%; strongly agree: 2/126, 1.6%; Figure 4; Multimedia Appendix 8).

Figure 4. Actually observed positive health care effects of digital health applications (DiHA).

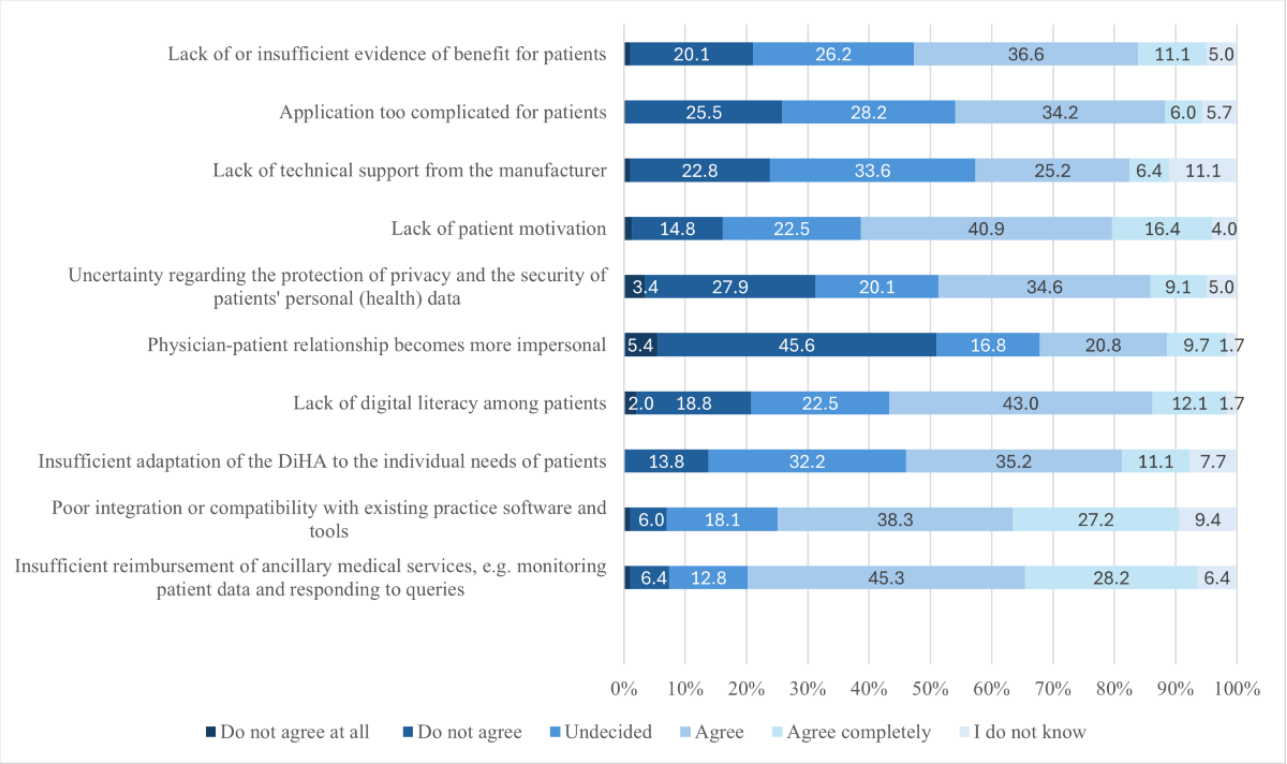


We found weak statistically significant correlations between the assessment of the individual positive health care effects already achieved and the intention to prescribe, except for improvements in access to hard-to-reach patient groups. A medium, statistically significant correlation was observed between the intention to prescribe and the improvement in QoL as a perceived positive health care effect (ρ=0.723; *P*=.002; Multimedia Appendix 9).

Barriers

A total of 298 HCPs evaluated the potential problems and barriers to prescribing HM-DiHA. Three-quarters of respondents identified the greatest barrier as the insufficient reimbursement for ancillary medical services, such as monitoring patient data or responding to queries (agree: 135/298, 45.3%; strongly agree: 84/298, 28.2%). More than half of the HCPs also cited barriers related to poor integration or compatibility with existing practice software, as well as a lack of digital affinity and motivation among patients. The lowest barriers were associated with an increasingly impersonal physician-patient relationship (agree: 62/298, 20.8%; strongly agree: 29/298, 9.7%) and a lack of technical support from the manufacturer (agree: 75/298, 25.2%; strongly agree: 19/298, 6.4%; Figure 5; Multimedia Appendix 10).

Figure 5. Barriers to digital health applications (DiHA) prescription.



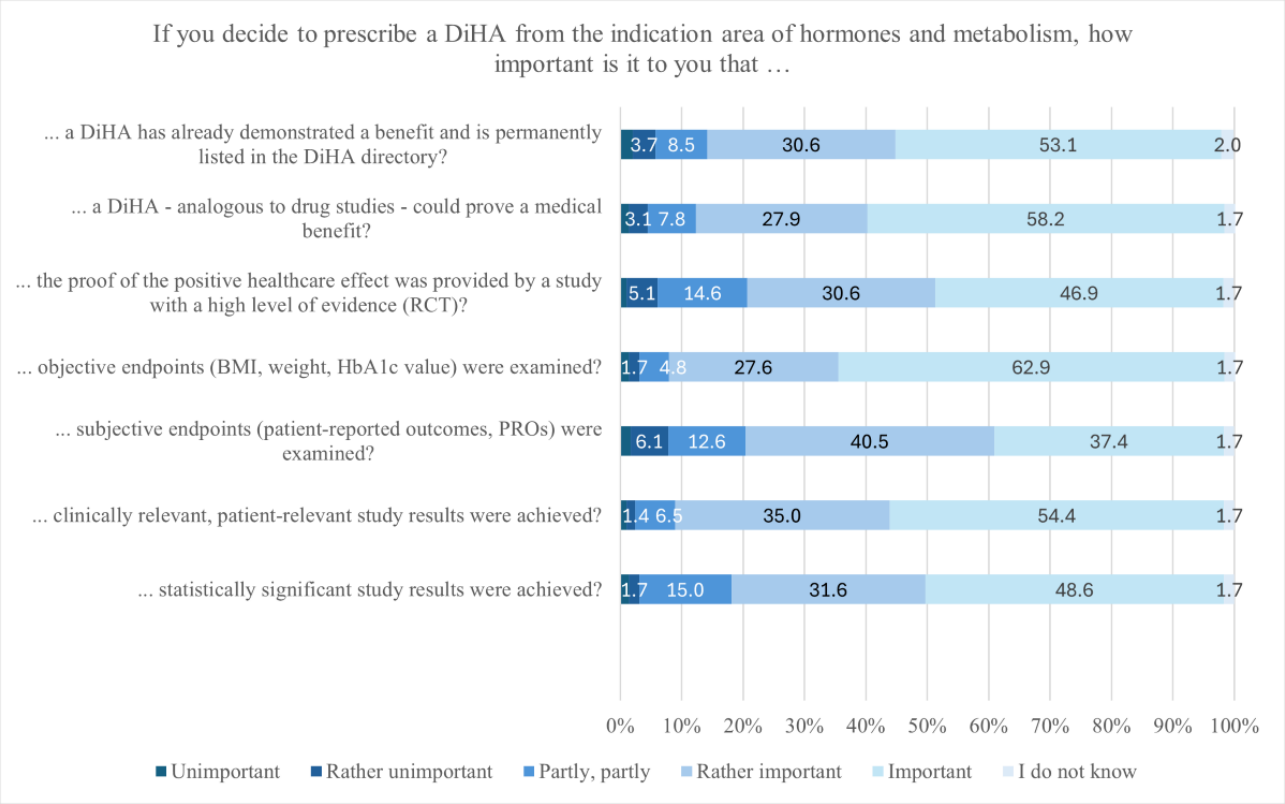
Weak, statistically significant correlations were found between the assessment of barriers—such as a lack of or insufficient proof of benefit for patients (*V*=0.247; *P*=.003), a lack of technical support from the manufacturer (*V*=0.195; *P*=.046), poor integration or compatibility with existing practice software (*V*=0.241; *P*=.004), and insufficient reimbursement for ancillary medical services (*V*=0.223; *P*=.001)—and the general DiHA prescription. No statistically significant correlations were observed regarding potential barriers and the prescription of HM-DiHA, nor with prescription frequency, except for the insufficient adaptation of DiHA to the individual needs of patients (ρ=–0.382; *P*<.001). However, numerous statistically significant, negative correlations were identified between the perceived barriers and the willingness to prescribe within the next 12 months (Multimedia Appendix 11).

There was also an option to answer an open-ended question to provide additional problems and barriers, which 43 HCPs used. Financial aspects were the most frequently mentioned here. On one hand, HCPs considered their own remuneration for prescribing DiHA insufficient; on the other hand, they also found the costs of DiHA too high. Evidence was also cited as a barrier, with respondents pointing to the lack of long-term studies and the absence of patient feedback regarding the effectiveness of DiHA. The acceptance and motivation of patients were another challenge. Despite initial motivation, many patients discontinue DiHA use. Additionally, there is often a lack of basic acceptance of digital interventions or sufficient digital competence (particularly among older patients), which is a fundamental prerequisite for their use. Other challenges mentioned were related to the HCPs' digital affinity and their knowledge of DiHA. This includes a lack of awareness about the existence of these applications, insufficient information about DiHA services and how they work, their suitability for specific patient groups, and a lack of training opportunities.

DiHA Evidence

A total of 294 HCPs answered the questions to assess the relevance of the DiHA evidence. With regard to the evidence of DiHA, the HCPs consider it most important that objective end points such as BMI, weight, or HbA_{1c} value were examined (rather important: 81/294, 27.6%; important: 185/294, 62.9%), that clinically relevant, patient-relevant study results were achieved (rather important: 103/294, 35.0%; important: 160/294, 54.4%), that the DiHA—analogous to drug trials—was able to demonstrate a medical benefit (rather important: 82/294, 27.9%; important: 171/294, 58.2%) and that the DiHA has already demonstrated a benefit and is permanently listed in the DiHA directory (rather important: 90/294, 30.6%; important: 156/294, 53.1%; Figure 6).

Figure 6. Digital health applications (DiHA) evidence.



Discussion

Overview

To date, no research in scientific literature explores the prescribing experiences, as well as the potential and challenges of HM-DiHA, from the perspective of HCPs. Therefore, we conducted a quantitative data collection (n=350) in collaboration with the GDA to address this research gap. The primary aim of the study was to explore the experiences HCPs have had with prescribing DiHA in general, and specifically in the area of hormones and metabolism, as well as to assess their prescription intention, and to identify the positive health care effects and barriers associated with DiHA use and prescription. Based on the findings, recommendations for action were to be developed to enhance the acceptance of DiHA among HCPs.

Experience With Previous DiHA Prescription

Our results show that DiHA have not fully arrived in statutory standard care yet—although they can be prescribed since 2020—as more than half of the respondents (187/350, 53.4%) have not yet prescribed DiHA in general. It should be noted that the survey participants likely represent a sample generally more open to digitalization, which could indicate an even lower adoption rate in the broader population. However, compared to previous surveys, this figure has risen significantly. In these earlier surveys, the proportion of participants who had already prescribed a DiHA was 10% (103/1299; period: 2020-2021) [15], 14% (536/3829; time point: 2022) [16] and 59.4% (58/97; time point: 2023) [18]. However, the results of Brecher et al [18] also show that half of those GPs who had already prescribed a DiHA only did so less than once a month (39/97, 49%). This proportion is comparable to that in this study, which was 54% (75/139). Overall, the scientific literature shows that there is still a high degree of skepticism on the part of HCPs regarding DiHA [16,25].

Wangler and Jansky [16] showed that HCPs in urban settings (21%) were statistically significantly more likely to prescribe DiHA compared to HCPs who worked in rural areas (5%). This study also showed that there was a statistically significant correlation between the general prescription of a DiHA and the size of the municipality or city. Of the HCPs who had already prescribed a DiHA, 26.5% (36/136) worked in a municipality or city with 5000-20,000 or more than 500,000 inhabitants. Overall, DiHA prescription is considered suitable for all age groups, with people aged 26-35 years and 36-45 years being cited most frequently. Dahlhausen et al [15] also found that 40.7% (527/1295) of respondents would primarily prescribe DiHA to younger patients. Some HCPs have identified a lack of DiHA options for children and adolescents, suggesting that future development should focus on their specific needs and requirements. Studies highlight the potential of digital health interventions in treating children and adolescents with obesity and DM and advocate for their integration into existing treatment approaches [26,27]. Among HCPs who prescribed a DiHA in general, those who used manufacturer access were nearly 12 times more likely to have prescribed a DiHA, highlighting that targeted provision of such access can enhance DiHA acceptance.

Prescription Intention

Although our results showed a positive correlation between the numerous positive health care effects already perceived and the intention to prescribe, 42.9% (149/348) of respondents did not intend to prescribe DiHA in the next 12 months. This represents an increase compared to previous surveys, in which 37% (19/51;

time point: 2020) [14], 30.3% (393/1299; time period: 2020-2021) [15], 13% (498/3829; time point: 2022) [16], and 38.5% (37/97; time point: 2023) [18] reported similar intentions. Differences in study periods, survey methodologies, and target populations may have contributed to this variance. The proportion of respondents who remain uncertain about prescribing DiHA (44/348, 12.6%) has slightly decreased compared to Dahlhausen et al (259/1299, 19.9%) [15].

These results suggest that further research is needed to identify key factors influencing DiHA acceptance and, in turn, to develop measures that enhance acceptance among HCPs in the field of hormones and metabolism. To sustainably implement DiHA in SHI standard care, the GDA has already outlined initial approaches aimed at making HM-DiHA an integral part of existing and future disease management programs [28]. Obesity-related DiHA are already included in the medical guideline on "Prevention and Treatment of Obesity" [29]. However, integrating diabetes-related DiHA into current guidelines has not occurred yet. DiHA with proven benefits specifically for diabetes should be more firmly embedded in the treatment process through integration into medical guidelines, allowing both HCPs and patients to recognize their direct added value. This recommendation is further supported by Wangler and Jansky's [30] findings. They showed that diabetes specialists would generally be more willing to include health apps in patient care (26% much more willing; 54% somewhat more willing) if national care guidelines for diabetes specifically addressed the use of such interventions.

Apart from the activities within the framework of SHI and the working model, the sociodemographic characteristics of HCPs did not correlate with the intention to prescribe. There is also disagreement within the scientific literature regarding the influence of sociodemographic characteristics, such as age and gender, on the acceptance of digital interventions. Some studies have identified these factors as determinants [31], while others have not [32]. However, the working model impacts the intention to prescribe, with the majority of HCPs working in hospitals being either rather unlikely (20/54, 37%) or very unlikely (27/67, 40.3%) to prescribe DiHA. Until March 2022, DiHA were only used in outpatient care and could not be prescribed in the inpatient sector due to the lack of adjustments to the discharge management framework agreement. However, our results show that some HCPs still believe they cannot prescribe DiHA or choose not to do so due to the lack of follow-up care because they work in the clinical setting. In the inpatient setting, more information should be provided about the possibilities for prescribing DiHA and monitoring their progress.

Health Care Effects

Our results indicate that the majority of HCPs recognize both the potential positive health care effects that DiHA use can achieve and the positive health care effects that have already been realized through its use. A positive attitude toward DiHA and their acceptance among HCPs can only be achieved if HCPs perceive the potential benefits and advantages of DiHA use for their patients, further supported by findings that an HCP's willingness to prescribe is influenced by their own perception of the usefulness of a digital intervention [33]. Therefore, the positive health care effects and potential of DiHA should be transparently presented and made comparable through standardized quality indicators.

Most respondents have observed improvements in patient-reported outcomes (PROs) such as self-management, health literacy, QoL, adherence, patient autonomy, disease management, and weight management as a result of DiHA use. These findings emphasize the high relevance of PROs and align with those of Dahlhausen et al [15], who reported improvements in adherence (997/1294, 77%), health literacy (842/1294, 65%), and disease management (783/1294, 60.5%). Similarly, Wangler and Jansky [16] found improvements in compliance (95%), mobility (94%), health awareness and education (93%), and self-management (91%) as the most common health care effects. Other studies also support this perspective, confirming the improvement of these PROs through DiHA use [25].

The direct patient benefits of HM-DiHA have already been confirmed for those permanently listed in randomized controlled trials [12,13,34]. Initial trends regarding patient benefits have also been observed for DiHA that are in the preliminary listing stage [10,35]. For example, diabetes-related DiHA have demonstrated reduced HbA_{1c} levels, while obesity-related DiHA have shown decreased weight as part of their primary end points. However, our results reveal that only 41.3% (52/126) of HCPs observed a reduction in HbA_{1c} levels in their patients, and 23.8% (30/126) were unsure about this outcome. Regarding weight reduction, 53.2% (67/126) of HCPs observed a decrease, while 24.6% (31/126) were unsure. To strengthen HCPs' acceptance and confidence in the effectiveness of DiHA, manufacturers should demonstrate the sustained positive health care effects observed in studies with a high level of evidence. Additionally, they should consider including factors such as patient adherence and the intensity of use in their analyses.

Barriers

Overview

The high proportion of HCPs who are very unlikely (83/348, 23.9%) or somewhat unlikely (66/348, 19%) to prescribe HM-DiHA over the next 12 months or remain uncertain (44/348, 12.6%), can be attributed to the identified barriers. Given the established correlation between perceived barriers and the willingness to prescribe, it is crucial to specifically address and overcome these obstacles.

Barrier 1 (Financial Barrier): The Reimbursement for Accompanying Medical Services Related to the DiHA Prescription is Inadequate

Insufficient reimbursement for ancillary services, such as monitoring patient data or responding to queries, was identified as the main barrier by nearly three-quarters of respondents (219/298, 73.5%), compared with half of HCPs in a recent study [15]. Managing DiHA is seen as an additional responsibility without additional remuneration, so HPCs lack financial incentives to engage with DiHA.

Monetary factors, such as reimbursement and the costs of digital interventions, are among the strongest predictors of HCPs'

acceptance of digital interventions [32,33]. In particular, the long-term costs of the technology, as well as the costs of devices and applications [33]. Financial incentives for prescribing medical services [36], including those related to DiHA [15], can enhance HCPs' acceptance.

In Germany, the uniform value scale (Einheitlicher Bewertungsmaßstab), which includes a fee schedule, regulates DiHA remuneration. For permanently listed obesity DiHA, an additional flat rate of 64 points (€7.64; US $8.9) is available for monitoring and evaluation. However, this can only be billed once per treatment or illness. There are no fee schedule items for the permanently listed diabetes DiHA. The reimbursement for follow-up checks of provisionally listed DiHA is significantly lower than the standard flat rate for personal or video consultations in primary care [37]. This highlights a clear imbalance and may act as a barrier to DiHA adoption among HCPs.

Overall, our findings highlight the significant importance of reimbursing medical services related to DiHA prescriptions, with HCPs rating the current reimbursement as insufficient. To support the sustainable use of DiHA, clearly defined and adequately incentivized billing codes are needed—ones that realistically reflect the time and effort required by HCPs. Initial steps, such as the introduction of the uniform flat rate, mark a move in the right direction but remain insufficient for widespread implementation. A targeted refinement of the remuneration framework is therefore essential to embed DiHA not only technically, but also economically, into routine medical practice.

Barrier 2 (Financial Barrier): The Costs for a DiHA Are Too High

Another barrier raised by numerous HCPs was the perceived high cost of individual DiHA. In the first year after a DiHA is included in the DiHA directory, the manufacturer sets the actual price, regardless of whether the DiHA is listed preliminarily or permanently. To prevent arbitrary pricing by manufacturers, this price is capped by a maximum amount regulation, under which the National Association of SHI Funds (GKV-Spitzenverband) and the manufacturers' associations reach a joint framework agreement and set maximum amounts for 90-day DiHA usage. At the end of the first 12 months of a DiHA's listing, the reimbursement amount is negotiated between the GKV-Spitzenverband and the manufacturers. From January 1, 2026, at least 20% of the permanent remuneration amount for a DiHA will be performance-related, depending on its successful application. Success is determined through the "application-related performance measurement," the specific form of which is currently under discussion. To increase the acceptance of DiHA and the willingness to prescribe them, HCPs should be more involved in the entire pricing process for DiHA, as well as in the development of the application-related performance measurement.

Currently, the prices of HM-DiHA are generally higher compared to other indication areas [38], with prices ranging from €218 (US $255; zanadio) to €740 (US $867; Una Health for diabetes) [39]. A comparison over recent years reveals that the most significant increase—20%—occurred between 2021 and 2022. Since then, manufacturer prices have risen by an additional 11%. Overall, prices increased by 39%, equivalent to €162 (US $190), between 2020 and 2024 [40]. The cost of treatment significantly influences the prescribing decisions of HCPs [36], which is supported by the results of this study, with cheaper DiHA being prescribed more often compared to more expensive DiHA. However, prices should be structured in a way that does not overburden the health care system, which is financed on a solidarity basis, while still offering sufficient incentives for DiHA manufacturers to invest in the design and improvement of their products.

Powell and Torous [41] outlined a patient-centered framework for evaluating the economic value of the clinical benefits of DiHA. They proposed a model incorporating factors such as the country-specific value of quality-adjusted life years, engagement rates, and the app's health impact. By applying this model to 2 DiHA, the authors aim to provide country-specific estimates of their economic value, helping to guide future research in digital health. This approach can inform value-based payment models and clinical decision-making [41].

While health economic evaluations typically rely on incremental cost-effectiveness ratios—which represent the difference in cost between the intervention group and the control group divided by the difference in their effects—applying this approach in the DiHA context may be challenging, as German health authorities are fundamentally opposed to the use of general cost-effectiveness thresholds. Furthermore, deriving an appropriate incremental cost-effectiveness ratio through an alternative approach would necessitate the consideration of all relevant treatment options within the therapeutic area, making it a highly time-consuming process [42]. Gensorowsky et al [42] argued that a pragmatic approach to DiHA pricing could involve aligning their demonstrated benefits with the cost-effectiveness benchmarks of established SHI-covered treatments in the same therapeutic area. Therefore, a DiHA price would be considered appropriate if the cost per unit of benefit generated by the DiHA is equal to or lower than the cost per unit of benefit provided by the existing therapy [42]. Going forward, the evaluation of a DiHA should not only encompass its positive health outcomes but also include a comprehensive health economic analysis.

Barrier 3 (Acceptance and Motivation): There Is a Lack of Acceptance, Digital Affinity, and Patient Motivation

Additional barriers include patients' lack of motivation (171/298; 57.3%) and limited digital affinity (164/298; 55.1%). Many patients either do not accept DiHA or have insufficient digital literacy, particularly older individuals, as reflected by most respondents reporting that patients never requested a DiHA (26/139; 18.7%) or requested one less than once a month (76/139; 54.7%). Similarly, findings from Wangler and Jansky indicate that GPs are either never (41%) or only occasionally (18%) approached by patients regarding DiHA [16]. A survey of diabetes specialists further supports this, revealing that patients never (10%) or only rarely (38%) inquire about health apps for the prevention and management of T2DM [17].

Moreover, our survey revealed that patients frequently discontinue DiHA use, which may explain why follow-up

prescriptions are never issued (35/139; 25.2%) or issued less than once a month (72/139; 51.8%). A systematic review on the acceptance of mHealth interventions among HCPs also identified patients' digital affinity and motivation as key factors influencing acceptance [33]. These findings highlight the critical importance of educational efforts by HCPs to enhance patient engagement and sustained use of DiHA.

Barrier 4 (Evidence-Based Research and Data): There Is a Lack of Study Evidence and Long-Term Data

Nearly half of the HCPs surveyed (142/298, 47.7%) cited a lack or insufficiency of evidence for patient benefits as a barrier, consistent with Dahlhausen et al [15] (712/1298, 54.9%). At the same time, evidence related to DiHA—such as the inclusion of objective end points, the demonstration of clinically and patient-relevant outcomes, proof of medical benefit, and the resulting permanent listing in the DiHA registry—plays an important role for HCPs when deciding whether to prescribe an HM-DiHA. One contributing factor is the lack of patient feedback, which makes it difficult to assess the overall health care impact of DiHA. Additionally, the absence of long-term studies prevents a thorough evaluation of its sustained benefits. Given this context, it is crucial to ensure that all available evidence on the mechanisms of action and proven benefits of DiHA is easily accessible, standardized, and transparent—starting with the DiHA directory at the BfArM.

Even within the DiHA directory, information relevant to prescription decisions is often presented incompletely [43], despite being considered an essential tool for information and decision-making [14]. Wangler and Jansky [16] also highlight that the DiHA directory is viewed critically and lacks sufficient detail. Currently, the decision on whether a DiHA is granted preliminary or permanent inclusion in the directory is made solely by the BfArM. However, for future DiHA classified as risk class IIb—where the potential hazards are higher compared to lower-risk applications—HCPs should be involved in the approval process at an early stage. Furthermore, manufacturers should continue evaluating the efficacy of their DiHA even after permanent listing in the directory to address the need for more long-term data.

In accordance with the application-related performance measurement, patient adherence and the usage intensity of each application should also be considered [44]. It is important to acknowledge that adherence data, while offering an indication of users' engagement with the DiHA, do not necessarily capture the positive health care effect achieved. Although higher usage frequency may be associated with improvements in health status, it does not guarantee that the intended clinical end points are met. Consequently, for the design of application-related performance measurement, adherence should be interpreted only as a limited proxy for effectiveness. A more robust assessment requires combining adherence metrics with direct clinical end points, PROs, or other objective health indicators.

In addition, increasing the involvement of HCPs in the development of DiHA could help mitigate feelings of exclusion and disconnection from what is often perceived as a nonmedically driven development process [45]. Moreover, participatory technology development—aligned with a user-centered design approach—has the potential to enhance the long-term use and effectiveness of DiHA [46].

Barrier 5 (Knowledge and Information): There Is a Lack of Information and Knowledge Among HCPs

The lack of information and knowledge regarding the basic existence of DiHA, as well as their specific content, functionality, and suitability for certain patient groups, presents a significant challenge. In a recent study, the lack of information was identified as the greatest barrier (1135/1299, 87.4%) probably because the DiHA system was only recently introduced in Germany at the time of the study [15]. Even 3 years after the DiHA system was implemented, only 38.1% of GPs felt sufficiently informed to prescribe DiHA [18]. From an information management perspective, this persistent information gap illustrates deficits in the systematic provision, dissemination, and integration of knowledge relevant to DiHA. Effective information management in health care entails not only ensuring that accurate and relevant information is available, but also that it reaches the appropriate stakeholders at the right time and in an actionable format. The current situation reflects a lack of coordinated information governance and highlights the absence of an infrastructure that supports timely, user-oriented knowledge distribution for HCPs. To address future information deficits, there should be widespread education on the function and benefits of DiHA, as well as their prescription options. This educational effort should be as comprehensive as possible, involving the BfArM, the GKV-Spitzenverband, SHI physicians' associations, the state and federal medical associations, the manufacturers of DiHA, medical societies, and obesity and diabetes networks. In this context, these actors collectively form the core of a decentralized information management system whose collaboration is crucial to improving the flow, accessibility, and quality of information in the German DiHA system.

The role of medical associations should be emphasized here, as a survey of diabetes specialists revealed that 73% of respondents rely on the GDA for information on the evaluation and recommendation of mobile health apps [17]. The current body of research indicates that training can help fully leverage the potential of digital interventions. HCPs often fail to use the full range of features offered by digital interventions simply because they are unfamiliar with them [47]. Additionally, GPs would appreciate both written and in-person training on DiHA, along with support for its implementation in daily practice [14]. For this reason, training programs tailored to the target audience should be developed and made accessible. In this context, incentives for training—such as certification for continuing medical education—could be beneficial [15,16].

Limitations

The authors developed the online questionnaire based on preliminary work and a literature review. It was essential to develop a dedicated questionnaire for the DiHA survey targeting HCPs in the field of hormones and metabolism, as this specialty faces specific challenges and requirements in the use of DiHA. A tailored questionnaire allowed for a focused exploration of the unique needs and experiences of these HCPs, particularly

regarding the integration of DiHA into the treatment of complex hormonal and metabolic conditions.

While the existing scientific evidence was thoroughly identified, no systematic literature search was conducted. Additionally, the questionnaire underwent multistage pretesting, which helped to enhance its validity. An online survey offers an advantage through quick, contactless collection of large volume of data, ensuring anonymity and minimizing social desirability bias. However, a drawback is that any queries could not be addressed directly during the survey; instead, they could only be answered via email after the survey was completed.

The questionnaire link was distributed via email to all HCPs registered with the GDA. This one-sided recruitment method introduces a potential bias, which should be considered when evaluating the representativeness of the results. However, it allowed for nationwide distribution of the survey in Germany, thus increasing the relevance of the findings for the country. The distribution of respondents was balanced, with 46.6% (163/350) having already been prescribed a DiHA and 53.4% (187/350) not having done so yet. It cannot be ruled out that HCPs with a stronger interest in digital interventions were more likely to participate in the survey, and thus, potential response bias should be considered when interpreting the results.

The response rate of 5.8% (350/6,035) in this study is lower than in previous studies on physicians' attitudes toward DiHA without a specific indication focus (6.4% [14]; 7% [15]; 28% [16]; 31% [17]; 48.8% [18]). A total of 78 cases had to be excluded due to missing data. Additionally, further missing values were identified during the subsequent questionnaire processing, resulting in a final sample of 290 HCPs who completed the survey in full. Given this response rate, the survey cannot claim to be fully representative. Before data analysis, the missing values were examined for patterns, but no correlations were found. Imputation procedures were deliberately avoided, as the missing values were not random but rather due to the deliberate nonresponse of the HCPs.

A major limitation in data analysis arose from the sample size, which reduced the statistical power of the tests and increased the risk of a type II error (beta error). This means that potentially existing effects within the population may not be detected [23]. To address this, we used binary logistic regression analysis to investigate potential causal relationships. This approach has relatively few requirements (eg, for each group formed by categorical predictors, n≥25; and independent variables must not be highly correlated, $r>0.7$) and is comparatively insensitive to outliers [23]. The omnibus test of the model coefficients confirmed the significance of the regression model and its overall performance. The Cox and Snell R^2 of 0.371 and Nagelkerke R^2 of 0.497 indicated adequate model quality. Additionally, the Hosmer-Lemeshow test ($\chi^2_8=3.713$; $P=.88$) revealed only minor deviations between the predicted and observed values, which was also confirmed by the overall prediction accuracy (predicted vs observed) of 80.3%. Nevertheless, the CIs are very wide in some cases, which could be due to the sample size. If the number of observations is low, particularly in one category of the dependent variable, this can result in less reliable CIs and, consequently, more uncertain estimates [23].

Conclusions

For more than 5 years, DiHA have been driving innovation in patient care, offering new solutions for both patients and HCPs while helping to bridge existing gaps in care. Physicians play a crucial role in integrating DiHA into standard care. The aim of our study was to gather insights from HCPs regarding their experiences with prescribing DiHA, their prescribing intentions, as well as the opportunities and challenges associated with HM-DiHA. Our results indicate that DiHA have not yet been fully integrated into standard care, highlighting the need for targeted measures to facilitate their adoption in the coming years. Based on our findings, we provide key recommendations for health policymakers, DiHA manufacturers, and DiHA prescribers. While these recommendations are strong and actionable, it is important to note that the findings are exploratory, cross-sectional, and based on a limited sample size, and should therefore be interpreted with appropriate caution.

1. Development of DiHA tailored to children and adolescents' specific needs and requirements.
2. Targeted provision of DiHA manufacturer access.
3. Strengthening the integration of DiHA into medical guidelines would help anchor it more firmly within the treatment process.
4. Information on the options for prescribing DiHA and monitoring their progress in inpatient settings.
5. Transparent presentation of positive health care effects and development of standardized quality indicators for their comparability.
6. Adequate compensation for accompanying medical services.
7. Inclusion of HCPs in the entire pricing process for DiHA, and the design of application-based performance measurement.
8. Health economic evaluation of DiHA.
9. Early involvement of HCPs in the approval process, particularly for DiHA with risk class IIb.
10. Ongoing effectiveness evaluation, even after permanent inclusion in the directory.
11. Comprehensive information and certified training on the functions, benefits, and prescription options of DiHA.

Furthermore, indication-specific research is needed to validate the findings and adapt them to other medical conditions. Future studies could also benefit from exploring the potential and challenges of DiHA from the perspective of HCPs across different countries and health care systems.

Acknowledgments

We express our sincere gratitude to the German Diabetes Association for their support in data collection. We also extend our heartfelt thanks to all participants for their valuable contributions to the survey. Furthermore, we thank Christina Mosetty for the linguistic revision of our manuscript.

Funding

This study was funded by the Open Access Publishing Fund of Leipzig University, which is supported by the German Research Foundation within the program Open Access Publication Funding. The funder had no role in the study design, data collection, analysis, interpretation of results, or the writing of the manuscript.

Data Availability

This study is based on primary data collected through a quantitative online survey, conducted with the support of the German Diabetes Association (Deutsche Diabetes Gesellschaft). The dataset generated and analyzed during this study is not publicly available due to participant privacy but may be made available from the corresponding author upon reasonable request.

Authors' Contributions

MM developed the concept and methodology of the manuscript, conducted the literature research, created the questionnaire, conducted the questionnaire pretests, analyzed and processed the data, wrote the text of the manuscript, and created the figures and tables. DMW organized and supported the recruitment of the survey participants via the German Diabetes Association. DMW, TW, TS, CMH, RH, MG, and DH supported the development of the questionnaire and the pretests and reviewed the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared

Multimedia Appendix 1
Survey instrument.
[DOCX File , 271 KB - formative_v9i1e77792_app1.docx]

Multimedia Appendix 2
Distribution of the main sociodemographic characteristics of the health care providers.
[DOCX File , 52 KB - formative_v9i1e77792_app2.docx]

Multimedia Appendix 3
Experience with previous digital health application prescriptions in relation to sociodemographic variables, n (%).
[DOCX File , 66 KB - formative_v9i1e77792_app3.docx]

Multimedia Appendix 4
Prescription intention in relation to sociodemographic variables, n (%).
[DOCX File , 64 KB - formative_v9i1e77792_app4.docx]

Multimedia Appendix 5
Reasons for the lack of digital health application prescription.
[PNG File , 100 KB - formative_v9i1e77792_app5.png]

Multimedia Appendix 6
Binary logistic regression analysis: general digital health application prescription and sociodemographic variables/digital affinity.
[DOCX File , 58 KB - formative_v9i1e77792_app6.docx]

Multimedia Appendix 7
Prescription frequency in relation to sociodemographic variables, n (%).
[DOCX File , 62 KB - formative_v9i1e77792_app7.docx]

Multimedia Appendix 8

Assessment of potential (N=325) and actual (n=126) healthcare effects, n, (%).
[DOCX File , 50 KB - formative_v9i1e77792_app8.docx]

Multimedia Appendix 9
Correlation between healthcare effects and prescription experience, prescription frequency and prescription intention.
[DOCX File , 50 KB - formative_v9i1e77792_app9.docx]

Multimedia Appendix 10
Assessment of barriers (n=298) for digital health application prescription, n (%).
[DOCX File , 47 KB - formative_v9i1e77792_app10.docx]

Multimedia Appendix 11
Correlation between barriers and prescription experience, prescription frequency and prescription intention.
[DOCX File , 47 KB - formative_v9i1e77792_app11.docx]

References

1. Huntsman DD, Bulaj G. Home environment as a therapeutic target for prevention and treatment of chronic diseases: delivering restorative living spaces, patient education and self-care by bridging biophilic design, e-commerce and digital health technologies. Int J Environ Res Public Health 2025;22(2):1-32 [FREE Full text] [doi: 10.3390/ijerph22020225] [Medline: 40003451]
2. Taylor ML, Thomas EE, Vitangcol K, Marx W, Campbell KL, Caffery LJ, et al. Digital health experiences reported in chronic disease management: an umbrella review of qualitative studies. J Telemed Telecare 2022;28(10):705-717. [doi: 10.1177/1357633X221119620] [Medline: 36346938]
3. Boima V, Doku A, Agyekum F, Tuglo LS, Agyemang C. Effectiveness of digital health interventions on blood pressure control, lifestyle behaviours and adherence to medication in patients with hypertension in low-income and middle-income countries: a systematic review and meta-analysis of randomised controlled trials. EClinicalMedicine 2024;69:1-18 [FREE Full text] [doi: 10.1016/j.eclinm.2024.102432] [Medline: 38333367]
4. Magliano D, Boyko E. IDF Diabetes Atlas 10th Edition. Brussels: International Diabetes Federation; 2021.
5. World Obesity Atlas 2024. World Obesity Federation. 2024. URL: https://data.worldobesity.org/publications/WOF-Obesity-Atlas-v7.pdf [accessed 2024-12-18]
6. Lin X, Li H. Obesity: epidemiology, pathophysiology, and therapeutics. Front Endocrinol (Lausanne) 2021;12:1-12 [FREE Full text] [doi: 10.3389/fendo.2021.706978] [Medline: 34552557]
7. Aschalew AY, Yitayal M, Minyihun A. Health-related quality of life and associated factors among patients with diabetes mellitus at the University of Gondar referral hospital. Health Qual Life Outcomes 2020;18(1):1-8 [FREE Full text] [doi: 10.1186/s12955-020-01311-5] [Medline: 32156282]
8. Sun H, Saeedi P, Karuranga S, Pinkepank M, Ogurtsova K, Duncan BB, et al. IDF diabetes atlas: global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. Diabetes Res Clin Pract 2022;183:1-23 [FREE Full text] [doi: 10.1016/j.diabres.2021.109119] [Medline: 34879977]
9. Zahedani AD, McLaughlin T, Veluvali A, Aghaeepour N, Hosseinian A, Agarwal S, et al. Digital health application integrating wearable data and behavioral patterns improves metabolic health. NPJ Digit Med 2023;6(1):1-15 [FREE Full text] [doi: 10.1038/s41746-023-00956-y] [Medline: 38001287]
10. Bretschneider MP, Roth L, Schwarz PEH. Effectiveness of a digital health application for the treatment of diabetes type II: a pilot study. J Clin Med 2023;12(19):1-12 [FREE Full text] [doi: 10.3390/jcm12196317] [Medline: 37834960]
11. Hermanns N, Ehrmann D, Finke-Groene K, Krichbaum M, Roos T, Haak T, et al. Use of smartphone application versus written titration charts for basal insulin titration in adults with type 2 diabetes and suboptimal glycaemic control (My Dose Coach): multicentre, open-label, parallel, randomised controlled trial. Lancet Reg Health Eur 2023;33:1-12 [FREE Full text] [doi: 10.1016/j.lanepe.2023.100702] [Medline: 37954005]
12. Gemesi K, Winkler S, Schmidt-Tesch S, Schederecker F, Hauner H, Holzapfel C. Efficacy of an app-based multimodal lifestyle intervention on body weight in persons with obesity: results from a randomized controlled trial. Int J Obes (Lond) 2024;48(1):118-126 [FREE Full text] [doi: 10.1038/s41366-023-01415-0] [Medline: 38017117]
13. Roth L, Ordnung M, Forkmann K, Mehl N, Horstmann A. A randomized-controlled trial to evaluate the app-based multimodal weight loss program zanadio for patients with obesity. Obesity (Silver Spring) 2023;31(5):1300-1310. [doi: 10.1002/oby.23744] [Medline: 37140392]
14. Radic M, Brinkmann C, Radic D, Waack M, Donner I, Stein L. Digitale gesundheitsanwendungen auf rezept: wie steht es um die aakzeptanz in der arzteschaft? Fraunhofer-Zentrum für Internationales Management und Wissensökonomie Imw. 2021. URL: https://www.imw.fraunhofer.de/content/dam/moez/de/documents/210303_Studie_Digitale%20Gesundheitsanwendungen%20auf%20Rezept_DiGAs.pdf [accessed 2025-12-18]

15. Dahlhausen F, Zinner M, Bieske L, Ehlers JP, Boehme P, Fehring L. Physicians' attitudes toward prescribable mHealth apps and implications for adoption in germany: mixed methods study. JMIR Mhealth Uhealth 2021;9(11):1-12 [FREE Full text] [doi: 10.2196/33012] [Medline: 34817385]
16. Wangler J, Jansky M. What potential and added value do DiGA offer for primary care?-Results of a survey of general practitioners in Germany. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2022 Dec;65(12):1334-1343 [FREE Full text] [doi: 10.1007/s00103-022-03608-w] [Medline: 36269336]
17. Wangler J, Jansky M. Attitudes and experiences of registered diabetes specialists in using health apps for managing type 2 diabetes: results from a mixed-methods study in Germany 2021/2022. Arch Public Health 2023;81(1):1-11. [doi: 10.1186/s13690-023-01051-0]
18. Brecher A, Herrgen J, Mueller T. DiGA in primary care and the influence of patient-related factors: results of a survey of family doctors. Stud Health Technol Inform 2024;313:55-61. [doi: 10.3233/SHTI240012] [Medline: 38682505]
19. Fakten zu diabetes - ein uberblick. Deutsche Diabetes Gesellschaft. 2025. URL: https://www.ddg.info/ddg-factsheet [accessed 2025-12-18]
20. AppQ: Gütekriterien-Kernset für mehr Qualitätstransparenz bei digitalen Gesundheitsanwendungen. Bertelsmann Stiftung. 2019. URL: https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/Studienbericht_AppQ_191028.pdf [accessed 2025-12-18]
21. Giebel GD, Speckemeier C, Abels C, Plescher F, Börchers K, Wasem J, et al. Problems and barriers related to the use of digital health applications: scoping review. J Med Internet Res 2023;25:1-16 [FREE Full text] [doi: 10.2196/43808] [Medline: 37171838]
22. Mäder M, Timpel P, Schönfelder T, Militzer-Horstmann C, Scheibe S, Heinrich R, et al. Evidence requirements of permanently listed digital health applications (DiGA) and their implementation in the German DiGA directory: an analysis. BMC Health Serv Res 2023;23(1):1-25 [FREE Full text] [doi: 10.1186/s12913-023-09287-w] [Medline: 37069592]
23. Döring N, Bortz J, Pöschl S, Werner C, Schermelleh-Engel K, Gerhard C. Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften. 5th ed. Berlin, Heidelberg: Springer Berlin Heidelberg; 2016.
24. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Hillsdale: Erlbaum; 1988.
25. Kreuzenbeck CCJ, Schneider BS, Brenner SX, Koerber F. Rapid review on the incentives of digital health apps for physicians and psychotherapists: a German perspective. Digit Health 2024;10:1-12 [FREE Full text] [doi: 10.1177/20552076241242781] [Medline: 38698827]
26. Pawellek S, Ziegeldorf A, Wulff H. Strategies and effects of digital interventions in overweight and obesity treatments in children and adolescents-a systematic review. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2022;65(5):624-634 [FREE Full text] [doi: 10.1007/s00103-022-03512-3] [Medline: 35320378]
27. Knox ECL, Quirk H, Glazebrook C, Randell T, Blake H. Impact of technology-based interventions for children and young people with type 1 diabetes on key diabetes self-management behaviours and prerequisites: a systematic review. BMC Endocr Disord 2019;19(1):1-14 [FREE Full text] [doi: 10.1186/s12902-018-0331-6] [Medline: 30630442]
28. Stellungnahme der Deutschen Diabetes Gesellschaft zum Entwurf eines Gesetzes zur Beschleunigung der Digitalisierung des Gesundheitswesens (Digital-Gesetz-DigiG). Deutsche Diabetes Gesellschaft. 2023. URL: https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/3_Downloads/Gesetze_und_Verordnungen/Stellungnahmen_WP20/DigiG/stellungnahme_ddg.pdf [accessed 2025-12-18]
29. German Obesity Association. Interdisziplinäre Leitlinie der Qualität S3 zur Prävention und Therapie der Adipositas. Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften. 2024. URL: https://register.awmf.org/assets/guidelines/050-001l_S3_Praevention-Therapie-Adipositas_2024-10.pdf [accessed 2025-12-18]
30. Wangler J, Jansky M. Bedeutung und einsatzpotenziale von gesundheits-apps in der diabetologischen Versorgung – ergebnisse einer befragungsstudie. Diabetologie und Stoffwechsel 2022;18(01):69-77. [doi: 10.1055/a-1859-6400]
31. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Quarterly 2003;27(3):425-478. [doi: 10.2307/30036540]
32. Jacob C, Sanchez-Vazquez A, Ivory C. Social, organizational, and technological factors impacting clinicians' adoption of mobile health tools: systematic literature review. JMIR Mhealth Uhealth 2020;8(2):1-30 [FREE Full text] [doi: 10.2196/15935] [Medline: 32130167]
33. Gagnon M, Ngangue P, Payne-Gagnon J, Desmartis M. m-Health adoption by healthcare professionals: a systematic review. J Am Med Inform Assoc 2016;23(1):212-220 [FREE Full text] [doi: 10.1093/jamia/ocv052] [Medline: 26078410]
34. Balzus L, Feiler M, Ebert D. Studienbericht: HelloBetter diabetes und depression. HelloBetter. 2021. URL: https://cdn.hellobetter.de/blog/2021/08/Studienbericht_Diabetes.pdf [accessed 2020-12-25]
35. Tamez-Pérez HE, Deceased Author, Cantú-Santos OM, Gutierrez-González D, González-Facio R, Romero-Ibarguengoitia ME. Effect of digital-tool-supported basal insulin titration algorithm in reaching glycemic control in patients with type 2 diabetes in mexico. J Diabetes Sci Technol 2022;16(6):1513-1520 [FREE Full text] [doi: 10.1177/19322968211034533] [Medline: 34323110]
36. Davari M, Khorasani E, Tigabu BM. Factors influencing prescribing decisions of physicians: a Review. Ethiop J Health Sci 2018;28(6):795-804 [FREE Full text] [doi: 10.4314/ejhs.v28i6.15] [Medline: 30607097]

37. Apps auf Rezept: Digitale Gesundheitsanwendungen: Hinweise zur Verordnung, Abrech-nung und Vergütung.. Kassenärztliche Bundesvereinigung. 2024. URL: https://www.kbv.de/praxis/verordnungen/diga [accessed 2025-12-18]
38. Timpel P, Mäder M, Scheibe S, Heinrich R, Schönfelder T. Verschreibung und vergütung im DiGA-dschungel. Blick in die versorgungspraxis. Kompendium Diabetes 2023:1-68 [FREE Full text]
39. DiGA-verzeichnis. Bundesinstitut für Arzneimittel und Medizinprodukte. 2025. URL: https://diga.bfarm.de/de/verzeichnis [accessed 2025-12-18]
40. Bericht des GKV-Spitzenverbandes über die Inanspruchnahme und Entwicklung der Versor-gung mit Digitalen Gesundheitsanwendungen (DiGA-Bericht). GKV-Spitzenverband. 2024. URL: https://www.gkv-spitzenver-band.de/media/dokumente/krankenversicherung_1/telematik/digitales/2024_DiGA-Bericht_final.pdf [accessed 2025-12-18]
41. Powell A, Torous J. A patient-centered framework for measuring the economic value of the clinical benefits of digital health apps: theoretical modeling. JMIR Ment Health 2020;7(10):1-5 [FREE Full text] [doi: 10.2196/18812] [Medline: 33124995]
42. Gensorowsky D, Witte J, Batram M, Greiner W. Market access and value-based pricing of digital health applications in Germany. Cost Eff Resour Alloc 2022;20(1):1-14 [FREE Full text] [doi: 10.1186/s12962-022-00359-y] [Medline: 35698135]
43. Heinrich R, Maeder M, Timpel P, Scheibe S, Schoenfelder T. Systematische Analyse des DiGA-Verzeichnisses am BfArM: Informationen für Verordnungsentscheidungen von digitalen Gesundheitsanwendungen sind oft unvollständig. 2022 Presented at: Deutscher Kongress für Versorgungsforschung (DKVF), German Medical Science GMS Publishing House; Oct 5-7; Dusseldorf URL: https://www.egms.de/static/en/meetings/dkvf2022/22dkvf239.shtml [doi: 10.3205/22dkvf239]
44. Gesetz zur beschleunigung der digitalisierung des gesundheitswesens. Bundesministerium für Gesundheit. 2024. URL: https://dip.bundestag.de/vorgang/gesetz-zur-beschleunigung-der-digitalisierung-des-gesundheitswesens-digital-gesetz-digig/303311 [accessed 2025-12-18]
45. Schroeder T, Haug M, Georgiou A, Seaman K, Gewald H. Evidence of how physicians and their patients adopt mHealth apps in germany: exploratory qualitative study. JMIR Mhealth Uhealth 2024;12:1-15 [FREE Full text] [doi: 10.2196/48345] [Medline: 38231550]
46. Harst L, Mäder M, Timpel P. Digital divide im kontext partizipativer technikentwicklung. In: Kernebeck S, Fischer F, editors. Partizipative Technikentwicklung im Sozial- und Gesundheitswesen. Interdisziplinäre Konzepte und Methoden. 1st Edition. Bern: Hogrefe; 2024.
47. Duhm J, Fleischmann R, Schmidt S, Hupperts H, Brandt SA. Mobile electronic medical records promote workflow: physicians' perspective from a survey. JMIR Mhealth Uhealth 2016;4(2):1-9 [FREE Full text] [doi: 10.2196/mhealth.5464] [Medline: 27268720]

Abbreviations

BfArM: Federal Institute for Drugs and Medical Devices [Bundesinstitut für Arzneimittel und Medizinprodukte]
DiHA: digital health applications
DM: diabetes mellitus
GDA: German Diabetes Association
GP: general practitioner
HbA1c: glycated hemoglobin
HCP: health care provider
HM-DiHA: hormones and metabolism digital health applications
ICD-10: International Statistical Classification of Diseases, Tenth Revision
OR: odds ratio
PRO: patient-reported outcome
QoL: quality of life
SHI: statutory health insurance
T2DM: type 2 diabetes mellitus

Edited by A Schwartz, M Balcarras; submitted 24.05.25; peer-reviewed by M Aasim, F Schliess; comments to author 03.12.25; accepted 08.12.25; published 31.12.25.

Please cite as:
Mäder M, Müller-Wieland D, Wiesner T, Schoenfelder T, Militzer-Horstmann C, Heinrich R, Geisler M, Häckl D
Prescribing Experiences, Potentials, and Challenges of Digital Health Applications in the Field of Hormones and Metabolism: Cross-Sectional Survey Study of Health Care Providers in Germany
JMIR Form Res 2025;9:e77792
URL: https://formative.jmir.org/2025/1/e77792
doi:10.2196/77792
PMID:

Health Care Utilization in Patients With Atopic Dermatitis Experiencing Topical Steroid Withdrawal: Observational Cross-Sectional Social Media Questionnaire Study

Alexander Shayesteh[1], MD, PhD; Maja af Klinteberg[1], MD, PhD; Sophie Vrang[2], MSc, RN; Gunnthorunn Sigurdardottir[3,4], MD, PhD; MariHelen Sandström Falk[5], MD, PhD; Mikael Alsterholm[6], MD, PhD

[1]Department of Public Health and Clinical Medicine, Dermatology and Venereology, Umeå University, Umeå, Sweden

[2]Swedish Asthma and Allergy Association, Stockholm, Sweden

[3]Department of Dermatology and Venereology in Östergötland, Linköping University, Linköping, Sweden

[4]Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Östergötland, Sweden

[5]Vasakliniken Dermatology Clinic, Gothenburg, Sweden

[6]Department of Dermatology and Venereology, Sahlgrenska Academy, Institute of Clinical Sciences, University of Gothenburg, Gröna stråket 16, Gothenburg, Sweden

Corresponding Author:
Mikael Alsterholm, MD, PhD
Department of Dermatology and Venereology, Sahlgrenska Academy, Institute of Clinical Sciences, University of Gothenburg, Gröna stråket 16, Gothenburg, Sweden

Abstract

Background: Topical steroid withdrawal (TSW) is a controversial skin condition among health care providers due to a lack of evidence, but it has an impactful and growing presence on social media. There are few previous reports of health care utilization for symptoms attributed to TSW.

Objective: This study aims to investigate health care utilization and requests as well as information sources for TSW among patients with atopic dermatitis (AD).

Methods: This observational cross-sectional study used a questionnaire aimed at adults with AD, experiencing symptoms they attribute to TSW. The questionnaire was posted as a link, free to share with others, in a Swedish TSW−themed Facebook group and remained accessible for 4 weeks. Descriptive statistics and topical text analysis on open-ended items were used to present and interpret the results.

Results: The participants (n=82) reported dermatologists (n=41, 50%), general practitioners (n=40, 49%), and practitioners of complementary and alternative medicine (CAM; n=32, 39%) as the most frequent health care contacts for TSW. However, among participants with ongoing symptoms attributed to TSW (n=68), ongoing health care contacts with general practitioners, dermatologists, and practitioners of CAM were reported by only 10% (n=7), 22% (n=15), and 13% (n=11), respectively. For symptoms attributed to AD, the frequencies of health care provider contacts were higher. Almost all participants had sought help from a general practitioner (n=81, 99%) or a dermatologist (n=76, 93%) at some point, and many had also consulted a practitioner of CAM (n=59, 72%). Among those with ongoing symptoms attributed to AD, 43% (n=26) had an ongoing contact with a dermatologist. Participant-requested help and support from health care providers included understanding and confirmation of TSW impairments (n=45, 56%), treatment of symptoms (n=26, 32%), and increased awareness and information about TSW from health care providers (n=21, 26%). The most common TSW information sources were Facebook (n=78, 96%), websites (n=75, 93%), and Instagram (n=45, 56%), but YouTube (n=11, 14%), podcasts (n=7, 10%), and TikTok (n=5, 6%) were also reported.

Conclusions: This study investigates health care utilization patterns related to TSW. The results indicate that the participants received insufficient support from health care providers for symptoms they attributed to TSW. The participants initiated and maintained health care provider contacts for symptoms attributed to AD to a greater extent than for TSW and sought information and support for TSW elsewhere. Targeted interventions to overcome this could be educational efforts for general practitioners and dermatologists about the current scientific knowledge of TSW as well as the TSW discourse on social media. In addition, health care providers need to engage and contribute to evidence-based content about TSW on relevant social media platforms to prevent the spread of misinformation about topical glucocorticoids.

(JMIR Form Res 2025;9:e85183) doi:10.2196/85183

KEYWORDS

atopic dermatitis; red skin syndrome; topical steroid addiction; topical steroid withdrawal; topical steroid withdrawal syndrome

Introduction

Background

Topical steroid withdrawal (TSW) is described as an adverse skin reaction following the tapering or discontinuation of potent topical glucocorticoids (TGCs) [1,2]. It has also been reported to occur during TGC use [3]. The skin becomes intensely red, with itching, stinging, or pain [4,5]. Large skin areas are affected with frequent involvement of the face, and specific signs such as “red sleeve” (involvement of the arms with sparing of the hands) and “elephant wrinkles” (thickened skin with reduced elasticity over joints) have been documented [6]. Although first described in the late 1960s, TSW is a concept that has attracted rapidly increasing attention and concern among patients in the era of social media [3,7-9]. Meanwhile, TSW remains controversial among health care providers as there is neither clear evidence to support the suggested mechanism of addiction followed by symptoms of withdrawal nor a clinical definition of the condition [7,10-12]. A suggestion for the management of symptoms attributed to TSW, including cessation of TGCs, psychosocial support, and systemic treatment with dupilumab, was just provided using the modified Delphi consensus method [13]. Recently, steroid-induced dysregulation in nicotinamide adenine nucleotide metabolism has been demonstrated in clinical samples and proposed as an explanation for the symptoms described in TSW [14]. The conclusions drawn from this small study have been criticized [15].

Some aspects of health care utilization for symptoms attributed to TSW have been described. In retrospective reviews of case notes for patients in the United Kingdom and Australia, complementary medical practices, psychology support, and online research were reported [5,16]. Qualitative studies have explored health-seeking behavior and interactions with medical professionals and found patient withdrawal from standard dermatology care and experiences of dismissal when voicing TSW concerns [17,18]. Social media is abundant with TSW content, and it is assumed that this serves as a major source of information and support for patients [9,19,20]. The TSW discourse on social media is mostly based on accounts from individuals who identify as sufferers of TSW, with little or no contribution from health care providers [8,9,21,22].

Prior Work

The lack of a definition makes the TSW population hard to identify and reach within the health care system. Therefore, we conducted a questionnaire study on Swedish social media aimed at individuals with atopic dermatitis (AD) who had experienced symptoms that they attributed to TSW. We have previously reported that the investigated cohort was young, predominantly female, mostly self-diagnosed with TSW, and that the skin signs and symptoms that they attributed to TSW caused a significant negative life impact [4].

Study Objectives

The aim of this study was to describe TSW-related health care utilization, requests, and information sources as reported by our previously characterized cohort. The results could help to define targets for intervention to improve the management of TSW.

Methods

Questionnaire Design and Distribution

A 47-item questionnaire was constructed in Swedish in SurveyMonkey (SurveyMonkey Inc), as described elsewhere [4]. Checklist 1 describes the study design and reporting with the CHERRIES (Checklist for Reporting Results of Internet E-Surveys) [23]. In short, the questionnaire was designed and modified in a stepwise fashion where items were consecutively evaluated by the authors, a focus group of dermatologists from the steering group of SwedAD (the Swedish nationwide registry for patients with AD receiving systemic pharmacotherapy), and a focus group of patients with AD recruited from the Swedish Asthma and Allergy Association. Questionnaire items were multiple-choice or open-ended. Multimedia Appendix 1 features the informed consent and the questionnaire, translated into English. Adaptive questioning was used. Figure S1 in Multimedia Appendix 2 shows the flowchart for questionnaire items.

It was not considered feasible or appropriate to recruit patients in a clinical setting as there are no acknowledged diagnostic criteria for TSW. In addition, it has been reported that individuals with symptoms they attribute to TSW can feel dismissed in health care settings [18]. Clinical recruitment could therefore have excluded a substantial portion of the target population or only captured those already engaged in dermatological care. To reach individuals with relevant experiences, the questionnaire was instead fielded in a Swedish private TSW−themed Facebook group. At the time, this was the largest Swedish online community for TSW known to the authors. The administrator of the Facebook group was contacted by the patient representative in the research group (SV) with a request to post a link to the questionnaire [24]. The post with the link was made on April 24, 2023, describing the aim of the study, that individuals aged 18 years or older with previous or ongoing AD combined with previous or ongoing symptoms attributed to TSW were eligible as respondents, and encouraging dissemination of the link through sharing on social media where relevant. Reposts were made on May 9, 2023, and May 17, 2023, as reminders. The questionnaire remained open for 4 consecutive weeks.

Inclusion Criteria

The inclusion criteria were age ≥18 years, previous or ongoing symptoms attributed to AD, and previous or ongoing symptoms attributed to TSW. The questionnaire was terminated if an answer did not meet the inclusion criteria.

Topical Text Analysis

A topical text analysis was performed on the answers to open-ended items as previously described [4,25]. Topical text analysis refers to the identification of recurring themes and keywords in qualitative responses. Briefly, AS and MaK independently identified keywords and topics, whereafter consensus was reached through discussion. The result was triangulated with the other authors to identify and mitigate biases and strengthen the interpretations.

Data Analysis

IBM SPSS Statistics for Windows (version 28.0.1.1 [15]) was used for descriptive statistics. The results are shown as the number and percentage observed per category. Any missing data are indicated by presenting the total number observed for each item in tables and figures.

Ethical Considerations

Ethical approval for the study, including the study protocol, was sought with a request for an advisory statement from the Swedish Ethical Review Authority. As the study did not involve any intervention or processing of identifiable personal data, it was not subject to formal ethical review under Swedish legislation [26]. The advisory statement from the Swedish Ethical Review Authority confirmed that there were no ethical objections to the study (application number Dnr 2023-00189-01).

When accessing the link to the questionnaire, the participants were presented with an informed consent statement that described the aim and method of the study, what was required from the participant, and the management of data. The informed consent statement translated from Swedish into English is available in Multimedia Appendix 1. The participants gave informed consent by proceeding to the questionnaire after taking part in the informed consent statement. No compensation was given to the participants. The participants were anonymous to the investigators and could not be identified or contacted.

Results

The questionnaire was entered by 98 individuals and completed by 82, hereafter referred to as participants. Figure S2 in Multimedia Appendix 2 shows the flowchart for the completion rate.

Figure 1 illustrates health care utilization for symptoms attributed to TSW and AD. The participants were asked to define any reported health care contact as “previous—not ongoing” or “ongoing.” The most frequently reported previous or ongoing health care contacts for symptoms attributed to TSW among all participants were dermatologists (n=41, 50%), general practitioners (n=40, 49%), and practitioners of complementary and alternative medicine (CAM; n=32, 39%).

Figure 2 shows that among participants with ongoing symptoms attributed to TSW (n=68), ongoing health care contacts with dermatologists, general practitioners, and practitioners of CAM were only reported by 22% (n=15), 10% (n=7), and 13% (n=11), respectively.

The participants could report none, any, or all of the item options “General practitioner,” “Dermatologist,” “Other specialist,” and “Practitioner of CAM.” Other specialists were reported as previous or ongoing contacts by 61% (n=50) and 21% (n=17) of participants for AD and TSW, respectively. This option could be specified in free text, but few participants chose to do so. For AD, pulmonologist or allergist (n=2), pediatrician or pediatric allergist (n=2), ophthalmologist (n=1), and nutritionist (n=1) were mentioned. For TSW, allergist (n=3) was mentioned.

No health care contact ever for symptoms attributed to TSW was reported by 46% (n=38). Despite ongoing symptoms attributed to TSW, 68% (46/68) reported no ongoing health care contact for those symptoms. All participants with any previous or ongoing health care contacts for symptoms attributed to TSW (54%, n=44) had consulted 2 or more of the categories “General practitioner,” “Dermatologist,” “Other specialist,” and “Practitioner of CAM.”

Figure 1. Frequencies of self-reported health care contacts for atopic dermatitis (AD) and topical steroid withdrawal (TSW). The data are from all participants in a Swedish observational, cross-sectional, social media questionnaire investigating TSW health care utilization, requests, and information sources in adults with AD combined with symptoms attributed to TSW. The link to the questionnaire was posted on April 24, May 9, and May 17, 2023, in a Swedish TSW−themed private Facebook group. Sharing of the link with others to reach as many individuals with relevant experiences as possible was encouraged in the posts. The questionnaire was open from April 24 to May 21, 2023. Participants could report multiple health care contacts. CAM: complementary and alternative medicine.

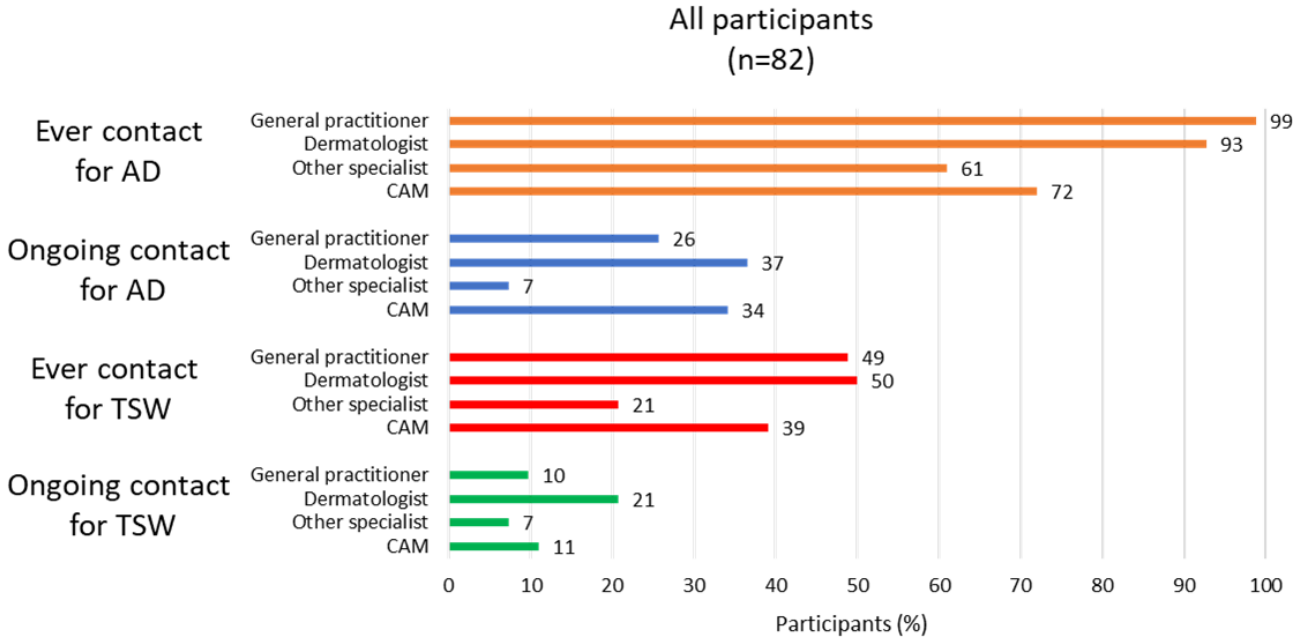


Figure 2. Frequencies of self-reported health care contacts for topical steroid withdrawal (TSW) among participants with ongoing symptoms attributed to TSW. The data are from a Swedish observational, cross-sectional, social media questionnaire investigating TSW health care utilization, requests, and information sources in adults with atopic dermatitis (AD) combined with symptoms attributed to TSW. The link to the questionnaire was posted on April 24, May 9, and May 17, 2023, in a Swedish TSW−themed private Facebook group. Sharing of the link with others to reach as many individuals with relevant experiences as possible was encouraged in the posts. The questionnaire was open from April 24 to May 24, 2023. As 53 participants reported both ongoing AD and TSW, the groups in Figures 2 and 3 overlap. Participants could report multiple health care contacts. CAM: complementary and alternative medicine.

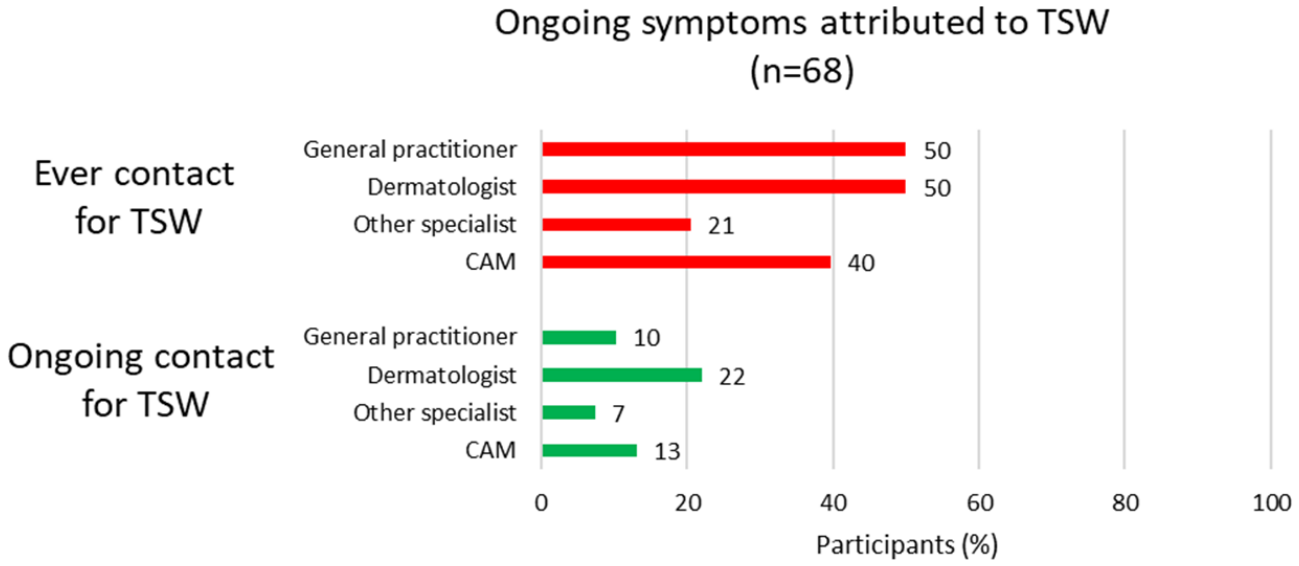


Figure 3. Frequencies of self-reported health care contacts for atopic dermatitis (AD) among participants with ongoing symptoms attributed to AD. The data are from a Swedish observational, cross-sectional, social media questionnaire investigating TSW health care utilization, requests, and information sources in adults with AD combined with symptoms attributed to TSW. The link to the questionnaire was posted on April 24, May 9, and May 17, 2023, in a Swedish TSW−themed private Facebook group. Sharing of the link with others to reach as many individuals with relevant experiences as possible was encouraged in the posts. The questionnaire was open from April 24 to May 21, 2023. As 53 participants reported both ongoing AD and topical steroid withdrawal, the groups in Figures 2 and 3 overlap. Participants could report multiple health care contacts. CAM: complementary and alternative medicine.

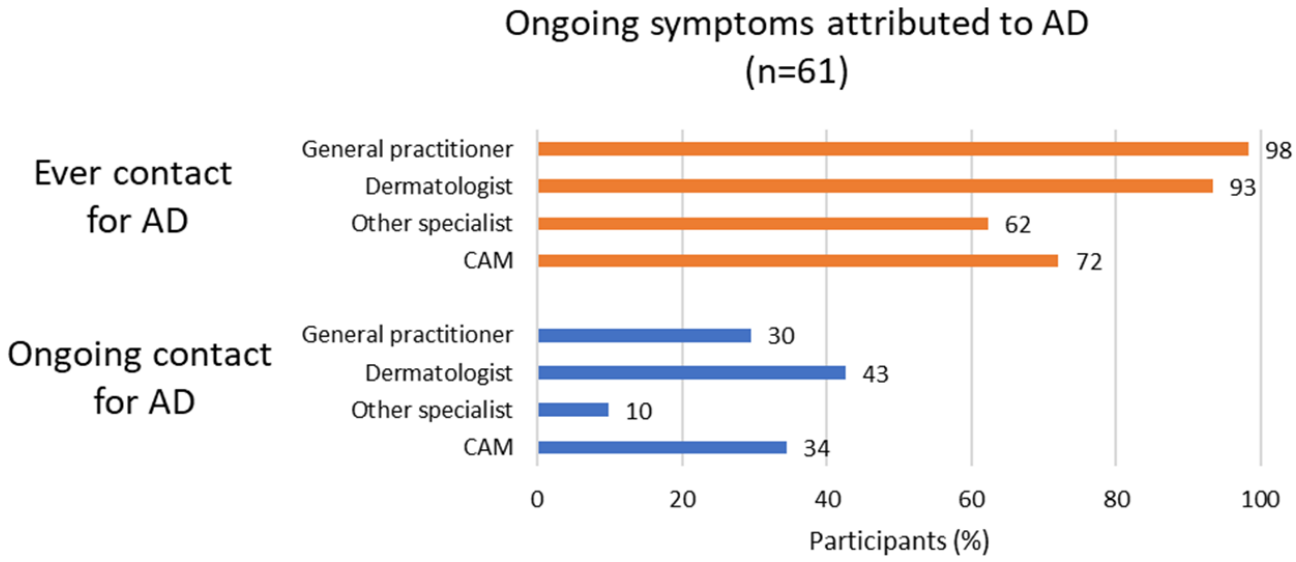


Table 1 describes investigations for symptoms attributed to TSW. Investigations beyond medical history and physical examination of the skin were unspecified blood tests (14/44, 32%), contact allergy tests (13/44, 30%, previously reported [4]), and skin biopsies 7% (3/44).

Figure 1 shows that both previous and ongoing health care contacts with general practitioners, dermatologists, other specialists, and practitioners of CAM were more common for symptoms attributed to AD than for symptoms attributed to TSW. Most participants had consulted a general practitioner (n=81, 99%) or a dermatologist (n=76, 93%) at some point, and many had also consulted a practitioner of CAM (n=59, 72%) or other specialists (n=50, 61%). Figure 3 illustrates that among participants with ongoing symptoms attributed to AD (n=61), ongoing health care contacts with dermatologists, general practitioners, and practitioners of CAM were reported by 43% (n=26), 30% (n=18), and 34% (n=21), respectively.

Table . Investigations for symptoms attributed to topical steroid withdrawal (TSW)[a].

Investigations	Participants, n (%)
Medical history on previous diseases and treatments	19 (43)
Physical examination of the affected skin areas by a doctor	19 (43)
Physical examination of all skin areas by a doctor	10 (23)
Skin biopsy	3 (7)
Blood test	14 (32)
Contact allergy test (applied on your back and checked after a few days)	13 (30)
No investigations have been performed	13 (30)
Other actions (please specify)	12 (27)[b]

[a]The data are from a Swedish observational, cross-sectional, social media questionnaire investigating TSW health care utilization, requests, and information sources in adults with atopic dermatitis combined with symptoms attributed to TSW. The link to the questionnaire was posted on April 24, May 9, and May 17, 2023, in a Swedish TSW−themed private Facebook group. Sharing of the link with others to reach as many individuals with relevant experiences as possible was encouraged in the posts. The questionnaire was open from April 24 to May 21, 2023. The table shows the frequencies of answers to the multiple-choice item “What investigations have been performed by your health care provider in regard to your TSW?” for the 44 participants who reported contact with a health care provider for symptoms attributed to TSW.

[b]The free-text answers given under the option “Other actions” were general comments about the participants’ experiences and did not specify any investigations.

Table 2 describes the participants’ requests for help and support for symptoms attributed to TSW. This was investigated with topical text analysis of the answers to the item “What kind of help and support would you like from the health care system?”. The most mentioned topics were “Support, understanding, and confirmation of TSW impairments” (n=45, 56%), “Medical help to alleviate symptoms” (n=26, 32%), and “Increased awareness and information about TSW from health care

providers" (n=21, 26%). "Medical investigation and a diagnosis" (n=8, 10%) and "Innovative treatments and research" (n=4, 5%) were less requested.

Table 3 presents the participants' self-reported TSW information sources. Social media platforms (Facebook: n=78, 96%; Instagram: n=45, 56%), and TSW-themed websites (n=75, 93%) were the most reported sources. YouTube and TikTok were not included as options but were mentioned under "Other options" by 14% (n=11) and 6% (n=5), respectively.

Table . Requested help and support for symptoms attributed to topical steroid withdrawal (TSW)[a].

Requested help and support	Participants, n (%)	Example of answers
Support, understanding, and confirmation of TSW impairments	45 (56)	"I need them (the healthcare providers) to understand and show empathy for my problems."
Medical help to alleviate symptoms	26 (32)	"We need a treatment programme for TSW."
Increased awareness and information about TSW from health care providers	21 (26)	"Someone educated in TSW should be on my side. As it stands now, I am told that TSW does not exist."
Paid sick leave	11 (14)	"I need paid sick leave. TSW patients should be able to rest in a stress-free environment."
Medical investigation and a diagnosis	8 (10)	"I want a proper medical investigation regarding causes, triggers, and cross-reactions in TSW."
Innovative treatments and research	4 (5)	"We need support for unconventional treatments and more research since some are affected by TSW while others are not."

[a]The data are from a Swedish observational, cross-sectional, social media questionnaire investigating TSW health care utilization, requests, and information sources in adults with atopic dermatitis combined with symptoms attributed to TSW. The link to the questionnaire was posted on April 24, May 9, and May 17, 2023, in a Swedish TSW−themed private Facebook group. Sharing of the link with others to reach as many individuals with relevant experiences as possible was encouraged in the posts. The questionnaire was open from April 24 to May 21, 2023. The table shows the topical text analysis of the answers to the item "What kind of help and support would you like from the health care system?". The frequencies of mentioned topics are presented, along with examples of answers. Participants (n=81) could mention more than 1 topic.

Table . Self-reported topical steroid withdrawal (TSW) information sources[a].

Information source	Participants, n (%)
Facebook groups	78 (96)
TSW-themed websites	75 (93)
Instagram	45 (56)
YouTube	11 (14)
Books or magazines or journals	11 (14)
Podcasts	8 (10)
TikTok	5 (6)
Blogs and online forums	4 (5)
TV or radio	N/A[b]
Twitter (now X)	N/A

[a]The data from a Swedish observational, cross-sectional, social media questionnaire investigating TSW health care utilization, requests, and information sources in adults with atopic dermatitis combined with symptoms attributed to TSW. The link to the questionnaire was posted on April 24, May 9, and May 17, 2023, in a Swedish TSW−themed private Facebook group. Sharing of the link with others to reach as many individuals with relevant experiences as possible was encouraged in the posts. The questionnaire was open from April 24 to May 21, 2023. The multiple-choice item "Which sources have you utilized for information about TSW?" was answered by 81 participants.

[b]N/A: not available.

Discussion

Principal Findings

This study investigates health care utilization for TSW, a primarily self-diagnosed condition where support and advice on management are offered by patient advocacy groups and users on social media platforms [27-29].

The frequencies of both previous and ongoing health care contacts were lower for symptoms attributed to TSW than for AD across all types of health care providers. Even with ongoing symptoms, it was less common to have a current health care

contact for TSW than for AD. We have previously described that the participants reported prominent and widespread skin signs, accompanied by pruritus, pain, and sleep disturbance, causing a considerable negative impact on social and intimate relations, anxiety, and feelings of depression [4]. It is striking and unfortunate that ongoing health care contacts were rare in relation to these severe manifestations. Ongoing AD treatment with systemic pharmacotherapy (dupilumab, methotrexate, upadacitinib) was reported by 22% (n=15) and phototherapy by 15% (n=12) of the participants, as previously presented [4].

The participants were more likely to initiate and maintain health care contacts for symptoms that they attributed to AD than for symptoms that they attributed to TSW. Presumably, this was partly explained by unmet needs in the management of TSW. The participants requested validation of their symptoms and suffering, targeted medical treatment, and information about TSW from health care providers. In addition, we have previously reported that reasons for not seeking medical help for TSW in this cohort were fear that the medical staff would not recognize the existence of TSW and a lack of knowledge of the condition [4]. Low confidence in health care providers' ability to offer the right treatment, experience or fear of being dismissed when expressing concerns regarding TGC safety, and being advised not to consult a health care provider were other reasons (data not shown).

The skin signs and symptoms attributed to TSW are similar to those described in insufficiently treated AD [30]. As the participants reported more ongoing health care contacts for AD than for TSW, even when they considered both conditions to be active, it can be speculated that some participants chose not to discuss the symptoms they attributed to TSW with their health care provider or that their concerns were ignored. TGCs are often part of the recommended treatment for AD. This can cause a dilemma for a patient with an unspoken attribution of symptoms to TSW, leading to nonadherence to treatment or ending of the health care contact.

For AD, ongoing contact with a practitioner of CAM and ongoing contact with a general practitioner were reported with a similar frequency, perhaps reflecting a tendency to seek health care with less focus on TGCs. On the other hand, ongoing contact with practitioners of CAM and ongoing contact with general practitioners were equally common for symptoms attributed to TSW as well, although at a lower level than for AD. Overall, 72% (n=59) had consulted a practitioner of CAM for AD and 39% (n=32) for symptoms attributed to TSW at some point. This is higher than what has been reported in an Australian cohort where only 6/55 (11%) had consulted a practitioner of CAM and, for the TSW-specific consultation, comparable to 9/19 (47%) reported from a clinic in the United Kingdom [5,16]. In Swedish surveys aimed at the general population, the use of any type of CAM, including self-prescribed and self-administered treatment, has been reported by 64% (318/500) to 71% (1089/1534) of the respondents [31,32]. Among those, 13% (65/500) to 33% (505/1534) reported consultation with a practitioner of CAM. The high levels of CAM consultations among the participants in this study should be interpreted in the context of the relatively common use of CAM services in Sweden but suggest that the participants were more inclined to consult with a practitioner of CAM than the general population. Variation in health care systems, cultural attitudes, and accessibility of CAM services could account for differences with comparable cohorts elsewhere.

Interestingly, only 10% (n=7) of the participants expressed a need for the investigation of the cause of their symptoms. It has been shown that patients tend to value and trust their own experiences of TGCs and may perceive health care providers' emphasis on the safety of TGCs as dismissive [18]. We have previously reported that 93% (76/82) of the participants in this study regarded TGCs as the principal trigger factor for their symptoms, with 33% (n=27) also identifying oral glucocorticoids as triggering [4]. Investigating other explanations, including but not limited to those related to TGCs, poses a challenge for health care providers. Undergoing skin biopsies and patch tests to address concerns of TSW was reported by a minority of the participants. Investigations are needed to understand symptoms attributed to TSW and should be offered when relevant. Still, it is likely that some patients may decline investigation and choose self-management strategies.

The reported health care utilization pattern in this study suggests that symptoms attributed to TSW were something that the participants dealt with on their own to a greater extent than symptoms attributed to AD. With limited scientific literature and almost no TSW information available from health care providers, patients turn to social media and TSW-dedicated websites for information and support [3,19,27]. Facebook was the most frequently reported source of TSW information, which likely reflects that the questionnaire was posted in a TSW-themed Facebook group. TSW-themed websites and Instagram were the 2 other major sources of information. It can be assumed that the participants' age (74% [61/82] were 18 - 39 years old) and the time of data collection (spring 2023) influenced the outcome. For instance, TikTok would probably have been reported by a higher number of participants in 2025. TikTok now has more than 1 billion monthly users, and its potential for sharing public health dermatology information has been recognized [33-35]. The TSW content on TikTok, where #topicalsteroidwithdrawal has more than 600 million views, was recently assessed with instruments for quality-grading of health care−related digital content [21]. The authors found that the 100 most viewed TSW videos were of low quality, lacked pertinent information, and were based on personal accounts with no health care provider engagement.

Strengths and Limitations

A strength of this study is that it was designed to reach as many individuals with relevant experiences as possible. Social media was chosen to field a free-to-share questionnaire because there are no established diagnostic criteria for TSW, making it hard to define and reach the intended population in the health care system [36,37]. The questionnaire was detailed, with a high completion rate, allowing the description of various aspects of TSW, some of which have been published [4]. For instance, the manifestations attributed to TSW and their life impact were

investigated, which was helpful in the assessment of the relevance of the health care contacts reported here.

Limitations of the study design are that all data were self-reported and that a response rate cannot be calculated. Recall bias and confirmation bias can affect data in controversial topics such as TSW. The results must be interpreted with caution as the external validity is unknown. The demography and attitudes of the participants are likely to reflect that a dedicated Facebook group was chosen to launch the questionnaire.

Multiple-choice items were used to investigate some aspects of TSW, which can introduce bias or loss of information. For instance, the item option "Other specialist" was not specified by most participants, but this information can be important as TSW-related complications requiring ophthalmologists have been reported [38]. No participants reported having seen a psychologist; however, this has been reported as high as 22% in the literature and noted in case reports [20,38,39]. The reason for this difference is not clear, but possible explanations are questionnaire design and the Swedish health care system. In retrospect, it might have been better to include specific items about psychological and psychiatric support. On the other hand, given the sensitive nature of mental health and the challenges in accessing individuals with TSW, including such items could have increased participant discomfort, the risk of social desirability bias, or non-response. In the Swedish public health care system, a referral (usually from a general practitioner) is required for outpatient psychiatrist and psychologist consultations and can be unavailable or expensive in private practice.

A larger number of open-ended items could have provided more data but at the risk of a lower completion rate. No rating of the severity of AD was included, making it difficult to assess if the frequencies of ongoing health care contacts were reasonable given the symptoms.

Conclusions

As research data on TSW begin to emerge, it is important for clinical researchers, health care providers, and patient advocacy groups to consider how and where the information should be put forward. As of now, patients navigate parallel realities. Social media is abundant with engaging and supportive TSW content of low quality, and sometimes even misinformation. Health care providers are unfamiliar with the TSW discourse and have difficulties in engaging in conversations about TSW. Interventions are needed to bridge this discrepancy [40]. For health care providers, it is a delicate task to increase engagement in the TSW discussion without inadvertently confirming theories that lack scientific support.

The pattern of health care contacts in this study suggests that education about the evidence base for TSW, including the social media discourse, should be aimed at general practitioners and dermatologists. Even though TSW is not confirmed as a distinct entity, health care providers must be equipped to discuss the topic and engage empathically with concerned patients.

The information sources reported here support that health care provider organizations should provide evidence-based information about TGCs and TSW on relevant social media platforms and other digital services to reach those with general concerns about TGCs and those who have symptoms that they attribute to TSW.

Acknowledgments

The authors are deeply grateful to the participants, who generously devoted time to sharing their experiences. The authors would also like to thank Professor Elisabeth Nylander and Professor Carl-Fredrik Wahlgren for invaluable support in the study design process and the members of Atopikerna, the Swedish Asthma and Allergy Association, who took part in the questionnaire design focus group.

Funding

The study was financed by Swedish governmental funding of clinical research (Avtal om Läkarutbildning och Forskning).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request. The questionnaire used to generate the data is available Multimedia Appendix 1.

Authors' Contributions

Conceptualization: AS (equal), GS (supporting), MA (lead), MaK (supporting), MHSF (supporting), SV (supporting)
Data curation: AS (supporting), MA (lead)
Formal analysis: AS (equal), MA (lead), MaK (equal)
Methodology: AS (lead), GS (supporting), MA (equal), MaK (supporting), MHSF (supporting), SV (supporting)
Project administration: AS (supporting), MA (lead)
Visualization: AS (lead), MA (equal)
Writing – original draft: AS (supporting), MA (lead)
Writing – review & editing: AS (lead), GS (supporting), MA (equal), MaK (supporting), MHSF (supporting), SV (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Informed consent and presentation of the questionnaire (translated from Swedish into English).
[PDF File, 157 KB - formative_v9i1e85183_app1.pdf]

Multimedia Appendix 2
Flowchart for questionnaire items and flowchart for questionnaire participation.
[PDF File, 78 KB - formative_v9i1e85183_app2.pdf]

Checklist 1
CHERRIES checklist.
[DOCX File, 29 KB - formative_v9i1e85183_app3.docx]

References

1. Hajar T, Leshem YA, Hanifin JM, et al. A systematic review of topical corticosteroid withdrawal ("steroid addiction") in patients with atopic dermatitis and other dermatoses. J Am Acad Dermatol 2015 Mar;72(3):541-549. [doi: 10.1016/j.jaad.2014.11.024] [Medline: 25592622]
2. Hwang J, Lio PA. Topical corticosteroid withdrawal ('steroid addiction'): an update of a systematic review. J Dermatolog Treat 2022 May;33(3):1293-1298. [doi: 10.1080/09546634.2021.1882659] [Medline: 33499686]
3. National Eczema Society, British Dermatological Nursing Group. Topical steroid withdrawal. : British Association of Dermatologists; 2024 URL: https://cdn.bad.org.uk/uploads/2024/02/22095550/Topical-Steroid-Withdrawal-Joint-Statement.pdf [accessed 2025-03-05]
4. Alsterholm M, Af Klinteberg M, Vrang S, Sigurdardottir G, Falk MS, Shayesteh A. Topical steroid withdrawal in atopic dermatitis: patient-reported characterization from a Swedish social media questionnaire. Acta Derm Venereol 2025 Jan 3;105:adv40187. [doi: 10.2340/actadv.v105.40187] [Medline: 39749386]
5. Brookes TSR, Barlow R, Mohandas P, Bewley A. Topical steroid withdrawal: an emerging clinical problem. Clin Exp Dermatol 2023 Aug 25;48(9):1007-1011. [doi: 10.1093/ced/llad161] [Medline: 37119282]
6. Sheary B, Harris MF. Cessation of long-term topical steroids in adult atopic dermatitis: a prospective cohort study. Dermatitis 2020;31(5):316-320. [doi: 10.1097/DER.0000000000000602] [Medline: 32404621]
7. Cotter C, Burton T, Proctor A, Moss C, Flohr C. Topical steroid withdrawal syndrome: time to bridge the gap. Br J Dermatol 2022 Nov;187(5):780-781. [doi: 10.1111/bjd.21770] [Medline: 35833942]
8. Finnegan P, Murphy M, O'Connor C. #corticophobia: a review on online misinformation related to topical steroids. Clin Exp Dermatol 2023 Feb 2;48(2):112-115. [doi: 10.1093/ced/llac019] [Medline: 36730502]
9. Nickles MA, Coale AT, Henderson WJA, Brown KE, Morrell DS, Nieman EL. Steroid phobia on social media platforms. Pediatr Dermatol 2023;40(3):479-482. [doi: 10.1111/pde.15269] [Medline: 36815409]
10. Awad A, Bell F, Ross G. Topical steroid withdrawal: a perspective of Australian dermatologists. Australas J Dermatol 2025 Nov;66(7):e433-e438. [doi: 10.1111/ajd.14580] [Medline: 40762418]
11. Spergel JM, Leung DYM. Topical steroid withdrawal syndrome: should we worry? Ann Allergy Asthma Immunol 2023 Jan;130(1):8. [doi: 10.1016/j.anai.2022.09.005] [Medline: 36596614]
12. Vroman F, Schimmel CS, Van der Rijst LP, et al. Topical steroid withdrawal: perspectives of Dutch healthcare professionals. Acta Derm Venereol 2025 Aug 12;105:adv44209. [doi: 10.2340/actadv.v105.44209] [Medline: 40792577]
13. Zemlok SK, Sandler M, Yu J. Initial consensus for the management of topical steroid withdrawal: a Delphi study. Dermatitis 2025 Oct 15. [doi: 10.1177/17103568251388635] [Medline: 41090229]
14. Shobnam N, Ratley G, Saksena S, et al. Topical steroid withdrawal is a targetable excess of mitochondrial NAD. J Invest Dermatol 2025 Aug;145(8):1953-1968. [doi: 10.1016/j.jid.2024.11.026] [Medline: 40088241]
15. Brown SJ, Burleigh A. Molecular and cellular mechanisms leading to topical steroid withdrawal syndrome remain unexplained and warrant further research. J Invest Dermatol 2025 Dec;145(12):3224-3225. [doi: 10.1016/j.jid.2025.05.026] [Medline: 40480485]
16. Sheary B. Steroid withdrawal effects following long-term topical corticosteroid use. Dermatitis 2018;29(4):213-218. [doi: 10.1097/DER.0000000000000387] [Medline: 29923852]
17. Sheary B, Tyson C, Harris MF. A qualitative study exploring the views and experiences of adults who decided to cease long-term topical steroid use. Dermatitis 2021;32(6):e131-e133. [doi: 10.1097/DER.0000000000000720] [Medline: 33654020]
18. Tan S, Phan P, Law JY, Choi E, Chandran NS. Qualitative analysis of topical corticosteroid concerns, topical steroid addiction and withdrawal in dermatological patients. BMJ Open 2022 Mar 16;12(3):e060867. [doi: 10.1136/bmjopen-2022-060867] [Medline: 35296492]

19. Bowe S, Masterson S, Murray G, Haugh I. Topical steroid withdrawal through the lens of social media. Clin Exp Dermatol 2022 Aug;47(8):1554-1557. [doi: 10.1111/ced.15194] [Medline: 35340034]
20. Orr N, Rogers M, Stein A, Thompson Coon J, Stein K. Reviewing the evidence base for topical steroid withdrawal syndrome in the research literature and social media platforms: an evidence gap map. J Med Internet Res 2024 Dec 6;26:e57687. [doi: 10.2196/57687] [Medline: 39642372]
21. Haddad F, Abou Shahla W, Saade D. Investigating topical steroid withdrawal videos on TikTok: cross-sectional analysis of the top 100 videos. JMIR Form Res 2024 Aug 29;8:e48389. [doi: 10.2196/48389] [Medline: 39208411]
22. O'Connor C, Murphy M. Scratching the surface: a review of online misinformation and conspiracy theories in atopic dermatitis. Clin Exp Dermatol 2021 Dec;46(8):1545-1547. [doi: 10.1111/ced.14679] [Medline: 33864398]
23. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34. [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
24. Whitaker C, Stevelink S, Fear N. The use of Facebook in recruiting participants for health research purposes: a systematic review. J Med Internet Res 2017 Aug 28;19(8):e290. [doi: 10.2196/jmir.7071] [Medline: 28851679]
25. Ferrario B, Stantcheva S. Eliciting people's first-order concerns: text analysis of open-ended survey questions. In: AEA Papers and Proceedings: American Economic Association; 2022, Vol. 112:163-169. [doi: 10.1257/pandp.20221071]
26. Act (2003:460) on ethical review of research involving humans [Web page in Swedish]. lagen.nu. URL: https://lagen.nu/2003:460 [accessed 2025-12-16]
27. ITSAN. URL: https://www.itsan.org [accessed 2025-03-05]
28. Taieb A. Topical steroid withdrawal: the challenging transition from patient self-diagnosis on social media to medical science and care. Acta Derm Venereol 2025 Apr 3;105:adv43455. [doi: 10.2340/actadv.v105.43455] [Medline: 40178221]
29. Marshall HF, Leung DYM, Lack G, et al. Topical steroid withdrawal and atopic dermatitis. Ann Allergy Asthma Immunol 2024 Apr;132(4):423-425. [doi: 10.1016/j.anai.2023.12.022] [Medline: 38142858]
30. Langan SM, Irvine AD, Weidinger S. Atopic dermatitis. Lancet 2020 Aug 1;396(10247):345-360. [doi: 10.1016/S0140-6736(20)31286-1] [Medline: 32738956]
31. Danell JAB, Wode K, Jong M, Kristoffersen AE, van der Werf E, Nordberg JH. Use of complementary and alternative medicine (CAM) in Sweden: a cross-sectional survey. Scand J Prim Health Care 2025 Sep;43(3):529-537. [doi: 10.1080/02813432.2025.2466180] [Medline: 39988756]
32. Wemrell M, Merlo J, Mulinari S, Hornborg AC. Two-thirds of survey respondents in southern Sweden used complementary or alternative medicine in 2015. Complement Med Res 2017;24(5):302-309. [doi: 10.1159/000464442] [Medline: 28535538]
33. Nguyen M, Youssef R, Kwon A, Chen R, Park JH. Dermatology on TikTok: analysis of content and creators. Int J Womens Dermatol 2021 Sep;7(4):488-489. [doi: 10.1016/j.ijwd.2021.04.006] [Medline: 34621965]
34. Villa-Ruiz C, Kassamali B, Mazori DR, Min M, Cobos G, LaChance A. Overview of TikTok's most viewed dermatologic content and assessment of its reliability. J Am Acad Dermatol 2021 Jul;85(1):273-274. [doi: 10.1016/j.jaad.2020.12.028] [Medline: 33359080]
35. Zheng DX, Mulligan KM, Scott JF. TikTok and dermatology: an opportunity for public health engagement. J Am Acad Dermatol 2021 Jul;85(1):e25-e26. [doi: 10.1016/j.jaad.2021.02.050] [Medline: 33639245]
36. Browne K. Snowball sampling: using social networks to research non‐heterosexual women. Int J Soc Res Methodol 2005 Feb;8(1):47-60. [doi: 10.1080/1364557032000081663]
37. Darko EM, Kleib M, Olson J. Social media use for research participant recruitment: integrative literature review. J Med Internet Res 2022 Aug 4;24(8):e38015. [doi: 10.2196/38015] [Medline: 35925655]
38. Sheary B. Topical steroid withdrawal: a case series of 10 children. Acta Derm Venereol 2019 May 1;99(6):551-556. [doi: 10.2340/00015555-3144] [Medline: 30734047]
39. Feschuk AM, Pratt ME. Topical steroid withdrawal syndrome in a mother and son: a case report. SAGE Open Med Case Rep 2023;11:2050313X231164268. [doi: 10.1177/2050313X231164268] [Medline: 37051261]
40. Guckian J, Hughes O, Nikookam Y, et al. Topical steroid withdrawal: self-diagnosis, unconscious bias and social media. Skin Health Dis 2025 Aug;5(4):281-288. [doi: 10.1093/skinhd/vzaf051] [Medline: 40755879]

Abbreviations

AD: atopic dermatitis
CAM: complementary and alternative medicine
CHERRIES: Checklist for Reporting Results of Internet E-Surveys
TGC: topical glucocorticoid
TSW: topical steroid withdrawal

Edited by A Mavragani; submitted 02.10.25; peer-reviewed by B Sheary; accepted 10.12.25; published 31.12.25.

Please cite as:
Shayesteh A, af Klinteberg M, Vrang S, Sigurdardottir G, Sandström Falk M, Alsterholm M
Health Care Utilization in Patients With Atopic Dermatitis Experiencing Topical Steroid Withdrawal: Observational Cross-Sectional Social Media Questionnaire Study
JMIR Form Res 2025;9:e85183
URL: https://formative.jmir.org/2025/1/e85183
doi:10.2196/85183

Original Paper

Characterization of Post-Viral Infection Behaviors Among Patients With Long COVID: Prospective, Observational, Longitudinal Cohort Analyses of Fitbit Data and Patient-Reported Outcomes

Tianmai M Zhang[1,2], MA; Sydney P Sharp[1], MPH; John D Scott[3], MSc, MD; Douglas Taren[4], PhD; Jane C Samaniego[1], MS; Elizabeth R Unger[5], MD, PhD; Jeanne Bertolli[5], MPH, PhD; Jin-Mann S Lin[5], PhD; Christian B Ramers[1,6,7,8], MPH, MD; Job G Godino[1,9,10], PhD

[1]Laura Rodriguez Research Institute, Family Health Centers of San Diego, San Diego, CA, United States

[2]Department of Biomedical Informatics and Medical Education, University of Washington, Seattle, WA, United States

[3]Department of Medicine, University of Washington, Seattle, WA, United States

[4]Department of Pediatrics, Nutrition Section, University of Colorado, Aurora, CO, United States

[5]National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, GA, United States

[6]School of Medicine, University of California, San Diego, La Jolla, CA, United States

[7]School of Public Health, San Diego State University, San Diego, CA, United States

[8]Global Hepatitis Program, Clinton Health Access Initiative, Boston, MA, United States

[9]Herbert Wertheim School of Public Health and Human Longevity Science, University of California, San Diego, La Jolla, CA, United States

[10]Exercise and Physical Activity Resource Center, University of California, San Diego, La Jolla, CA, United States

Corresponding Author:
Job G Godino, PhD
Laura Rodriguez Research Institute
Family Health Centers of San Diego
1750 5th Avenue
San Diego, CA, 92101
United States
Phone: 1 6195152344 ext 2344
Email: jobg@fhcsd.org

Abstract

Background: Long COVID encompasses a range of health problems that can be highly debilitating. While some research has relied on self-reported measures of symptoms and functioning, few studies have characterized symptoms in relation to behaviors and physiology measured objectively through wearable devices.

Objective: The primary aim of this study was to identify longitudinal patterns in physical activity, physiology, and patient-reported outcomes (PROs) among patients with long COVID at a Federally Qualified Health Center in the United States. The secondary aim was to identify meaningful subgroups or phenotypes within this cohort and examine how PROs and symptoms overlay with physical activity characteristics.

Methods: This was a prospective, observational, longitudinal cohort study recruiting a subset of low-income patients enrolled in the Long COVID and Fatiguing Illness Recovery Program. From March 2022 to May 2023, a total of 172 patients with long COVID or myalgic encephalomyelitis/chronic fatigue syndrome were given Fitbit Charge 5 (Fitbit Inc) devices and instructed to wear them continuously for up to a year. Patients completed PRO questionnaires (PROMIS-29 [Patient-Reported Outcomes Measurement Information System-29] and symptom questionnaires, etc) at baseline, 3, and 6 months. Inclusion in longitudinal analysis required at least 20 hours of valid wear data per day for a minimum of 7 days. The World Health Organization guideline on moderate to vigorous physical activity (MVPA) was used to differentiate MPVA-active and MVPA-inactive patients. Linear mixed effects regression was performed to assess longitudinal associations between physical activity levels and PROs.

Results: Among 172 patients, 80.2% (n=138) were female, 75.6% (n=130) were White, 45.3% (n=78) were unemployed, and 94.8% (n=163) had estimated annual income below US $50,000. Of these patients, 82 (47.7%) met valid wear criteria, providing 50.5 days of Fitbit data on average. At baseline, MVPA-inactive patients (n=41) reported statistically more severe problems regarding physical function, fatigue, and dyspnea than MVPA-active patients (n=41) on both continuous and categorical scales,

with $P<.05$ from both Student t tests (2-tailed) and chi-squared tests. Longitudinal analysis found that MVPA-inactive patients showed a decreased ability to participate in social roles (estimated group difference=–4.21 T-score points over 3 months, 95% CI –6.64 to –1.78, $P<.001$) and a higher intensity of sleep symptoms (estimated group difference=2.06 severity score points over 3 months, 95% CI 0.40 to 3.71, $P=.02$) over time.

Conclusions: This study showed that patients with long COVID could remain MVPA-active despite experiencing symptoms. These findings provide insights into the relationship between PROs, physical activity, and long COVID, which suggests the importance of considering individual activity profiles when tailoring treatment plans, especially in a low-income population. The findings of this study should be interpreted as hypothesis-generating, considering its exploratory nature and limitations, including high attrition rates and missing data.

(JMIR Form Res 2025;9:e77644) doi:10.2196/77644

KEYWORDS

long COVID; objective physical activity; phenotype; wearable device; Fitbit; post-viral infection behaviors; mobile phone

Introduction

COVID-19 has resulted in over 100 million infections and over 1 million deaths in the United States alone [1,2]. Patients who are infected with SARS-CoV-2 can experience a range of highly diverse clinical presentations, severities, and outcomes ranging from no symptoms to critical illness [3,4]. Some patients report lasting, new, or recurring symptoms and conditions more than 4 weeks after infection, commonly known as long COVID [4]. The scientific understanding of diagnoses, epidemiology, and phenotypes of long COVID is continually evolving; however, there is still much to be discovered and investigated [5]. Notably, there was not a consistent definition for long COVID until the National Academies of Sciences, Engineering, and Medicine recently defined it as an “infection-associated chronic condition (IACC) that occurs after SARS-CoV-2 infection and is present for at least 3 months as a continuous, relapsing and remitting, or progressive disease state that affects one or more organ systems” [6].

Long COVID encompasses a range of health problems and symptoms that may be highly debilitating and may arise secondary to detectable organ system damage, autoimmunity, and other mechanisms [6,7]. The recent Household Pulse survey estimated that among people who currently have long COVID, around 25% have significant activity-limiting symptoms [8]. In many patients, the etiopathogenesis is unclear [9]. As of September 2024, it was estimated that 18% of United States adults had experienced long COVID [8]. Typical symptoms associated with long COVID, such as cognitive impairment, fatigue, and post-exertional malaise (PEM), closely resemble those in patients diagnosed with myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) and other IACCs [10,11].

Increased risk of long COVID is associated with being female, experiencing socioeconomic deprivation, and being a member of an ethnic minority group [12,13]. Furthermore, individuals with lower incomes face greater stress, poorer access to health care, and environmental exposures compared to those with higher incomes [14]. These disparities highlight the need to study long COVID in medically underserved populations rather than in specialty clinics [15]. As California alone has about 5.9 million low-income residents, it is important to understand long COVID-related symptom experiences, physiological characteristics, and behavioral patterns in this population, which may support future improvements in the delivery of care.

Symptoms associated with long COVID are highly variable and are often unpredictable. An individual may experience day-to-day variation in symptoms and their severity [16,17]. Similarly, the way an individual responds to an event or stimulus is highly individualized and can often result in different behavioral outputs, depending on the day [18]. As the clinical presentation, severity, and outcomes of long COVID vary from patient to patient, it may prove beneficial to monitor the disease in a more personalized way, for example, through a wearable device.

Wearable devices that continuously track behavioral and physiological metrics such as steps, heart rate (HR), heart rate variability (HRV), temperature, physical activity, and sleep have become more accurate and accessible. It is estimated that one in three Americans uses a wearable device to help track their fitness and health [19]. Several studies and reviews have found that several wearable devices have demonstrated utility in infectious-disease surveillance, including early detection of COVID-19 and the monitoring of physiological signals pre and post infection [20-25].

As the wearable research landscape shifted from early detection to management of long COVID, several studies and reviews assessed a variety of wearables and approaches to identify persistent physiological clusters, support pacing or self-management, and track recovery trajectories [26,27]. However, the literature has a limited representation of low-income populations. Although many studies incorporated patient-reported outcomes (PROs), few used a longitudinal design to examine symptom trajectories alongside objective activity measures, and none did so in a Federally Qualified Health Center (FQHC) cohort.

To address this gap, we collected high-resolution longitudinal data from a subset of patients in the Long COVID and Fatiguing Illness Recovery Program (LC&FIRP) through Family Health Centers of San Diego (FHCSD) using the Fitbit Charge 5 (Fitbit Inc). This study sought to answer the following research questions: (1) What longitudinal characteristics do patients with long COVID have in terms of Fitbit-measured physical activity, physiology, and PROs? (2) Are there any meaningful subgroups or phenotypes within this patient cohort? (3) Are there

longitudinal associations between the identified phenotypes and the changes in PROs that could inform future studies and clinical management of long COVID?

Methods

Study Design and Setting

This was a prospective, observational, longitudinal cohort study involving a subset of patients enrolled in LC&FIRP. LC&FIRP was a three-year effectiveness-implementation hybrid study at an FQHC, FHCSD. FHCSD is one of the ten largest FQHC health systems in the nation, with the vast majority of its patients being low-income and members of a minority population [28]. The program used a technology-enabled multidisciplinary team-based care model focused on case-consultation and peer-to-peer discussion of emerging best practices (ie, teleECHO [Extension for Community Healthcare Outcomes]) to assist management of complex cases associated with long COVID, ME/CFS, and other IACCs [28]. This study was reported in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines [29]. The STROBE checklist of this study is available in Multimedia Appendix 1.

Aims

The primary aim of this study was to identify longitudinal patterns in physical activity, physiology, and PROs of patients with long COVID. The secondary aim was to identify meaningful subgroups or phenotypes within this study's cohort and examine the overlay of PROs and symptoms in relation to the activity characteristics of patients with long COVID.

Patients

Between March 2022 and May 2023, Fitbit devices for remote monitoring and symptom management were given to a subset (n=172) of the total enrolled LC&FIRP patient cohort (n=590). The sample size was based on 20% of LC&FIRP's initial recruitment target of 856. Using convenience sampling, participants were recruited when they agreed to complete their physical therapy visit or through a provider or physical therapist referral. Inclusion criteria included being aged 18 years or older, documented persistent symptoms and a decline in health-related quality of life consistent with long COVID based on patient report and clinical determination, ME/CFS or other IACCs, and having a smartphone. No additional measures were implemented to address potential sources of bias beyond standard study methods.

Eligible patients who completed their baseline survey and were willing to participate in the Fitbit substudy received Fitbit Charge 5 devices within 3 months. Participants also received a booklet explaining how to interpret Fitbit data to manage their symptoms using pacing. An activity diary was also included in the back of the booklet to encourage patients to record their activities and symptoms for a week straight. Patients were offered an incentive to submit the activity diary; however, only 4 patients submitted diaries, and therefore, these data are not included in the current analysis.

Patients completed PROs including the PROMIS-29 (Patient-Reported Outcomes Measurement Information System-29) version 2.1, the PROMIS (Patient-Reported Outcomes Measurement Information System) Short Form version 1.0 - Dyspnea Severity 10a and Applied Cognitive General Concerns 8a, the PHQ-2 (Patient Health Questionnaire-2), the GAD-7 (Generalized Anxiety Disorder-7), and the Centers for Disease Control and Prevention (CDC) ME/CFS Symptom Screener-Short Form version 1.2 at baseline, 3 months and 6 months [30-35]. They were completed over the phone and in English or Spanish, depending on the patient's preferred language. During follow-up, patients were only asked about symptoms they reported at baseline; however, patients had the opportunity to report any newly experienced symptoms during each follow-up. Data regarding HR, sleep, and physical activity were tracked by Fitbit devices in real time for up to a year and automatically uploaded to Fitabase (Small Steps Labs LLC), a Fitbit application programming interface partner that provided a platform for FHCSD to view clean data. After the 6-month survey date of the last patient in December 2023, Fitbit data was downloaded from Fitabase and preprocessed for analysis.

To minimize loss to follow-up, the research team contacted patients at least three times over 4 weeks to complete their surveys. The team also reviewed Fitabase weekly to confirm Fitbit data uploads. If a patient had not uploaded data for a week or more, researchers made three calls or text attempts over the next 4 weeks to remind them to wear the Fitbit and open the app to enable data upload.

Ethical Considerations

LC&FIRP was reviewed and approved by the Institutional Review Board of San Diego State University (HS-2021-0241). All methods were carried out in accordance with relevant guidelines and regulations, and present no more than minimal risk to patients. All patients included in this study completed FHCSD's Broad Consent, which includes a specific authorization for the use of deidentified health information for research purposes, including primary data collection and analyses. Any information obtained about patients during this study was deidentified and treated as strictly confidential to the full extent permitted by applicable law and in accordance with Health Insurance Portability and Accountability Act regulations. All data is deidentified and reported in aggregate. More information regarding data privacy and management can be found in the study protocol [28]. Patients kept the Fitbit Charge 5 devices they received after the current study as compensation for their participation in the Fitbit substudy, and a US $10 gift card was provided as an incentive for completing and returning the activity diary. No additional compensation was provided.

Data Preprocessing

Preprocessing of minute-level Fitbit data followed the recommendations of Wing et al [36]. Specifically, minutes with valid HR values were automatically considered valid wear. For minutes with missing HR data but with corresponding step counts, metabolic equivalent of task (MET), or intensity values above default (ie, step count above 0, MET above 1.0, or intensity above 1), the mean HR value from adjacent minutes

was used to fill in gaps [36]. Next, minutes with HR values below 45 or above 205 (0.23% and <0.001% of total observed minutes, respectively) were excluded due to their improbable occurrence. Additionally, periods where the HR value was unchanged for 11 or more consecutive minutes (1.39% of all observed minutes) were also excluded because they are unlikely to be physiological [36]. Subsequently, the minute-level time series data were aggregated into daily summaries to determine valid daily wear time and daily features of each participant. Summary measures for characterizing daily behaviors of patients were calculated from available Fitbit measurements. To retain longitudinal patterns in the Fitbit data, we did not conduct data imputation in consideration of the large proportion of missing data.

PRO scores were calculated according to their respective protocols. Severity scores of the 8 symptoms assessed by the CDC ME/CFS Symptom Screener were calculated according to the scoring algorithm of the CDC Symptom Inventory [34,35], which takes into account both the frequency and intensity of a symptom. Analyses of PROMIS scores were based on both raw scores and T-scores (ie, standard scores with a mean of 50 and SD of 10 in the US general population) calculated using the HealthMeasures Scoring Service online [37], while all other processes were performed using Python (version 3.8.15; Python Software Foundation). The Hawkins test suggested that the missingness pattern of the PRO data is consistent with a missing completely at random mechanism, with P>.10 for all PRO measures.

Patient Categorization

In this study, a valid wear day is defined as having at least 1200 minutes (ie, 20 hours) of valid Fitbit data. Patients are considered to have valid wear during this study if they have had at least 7 valid wear days among all available data.

Inspired by the findings from exploratory time series analysis that patients could be classified into 2 clusters by their moderate to vigorous physical activity (MVPA), we further categorized valid wear patients as MVPA-active or MVPA-inactive based on the World Health Organization (WHO) guideline of 150 MVPA minutes per week [38]. MVPA refers to the physical activity that is performed at >3 METs (ie, >3 times the intensity of rest) on an absolute scale [38]. Due to the discontinuous nature of valid wear days, we calculated average daily MVPA minutes of each patient between baseline and 6 months and categorized patients as MVPA-active or MVPA-inactive by whether their average daily MVPA reached 21.4 minutes.

Statistical Analysis

Statistical differences between subgroups of patients (eg, stratified by Fitbit wear time or MVPA) at different time points were examined using the chi-squared test, Student t test, or Wilcoxon test, according to data type and normality. Two-way ANOVA was used to examine group differences while adjusting for another factor of interest (eg, employment status).

For exploratory analysis of potential patient clusters, growth mixture modeling (GMM) [39] was applied on day-level time series of Fitbit measures.

Linear mixed effects (LME) [40] regression was used to model longitudinal associations between MVPA levels and PROs, using Fitbit data between baseline and 6 months. LME models of PROs included time (baseline, 3-month, and 6-month), MVPA levels, the interaction between time and MVPA levels, and baseline PRO scores, plus an individual intercept for each patient to account for random effects. The interaction term would reveal the effect of MVPA level over time. Models with different sets of independent variables or different random effect settings were evaluated with the Akaike Information Criterion to find the best model. Model validity was diagnosed using the Shapiro-Wilk test of residual normality and White Lagrange multiplier test of residual variance homogeneity. Sensitivity analysis was subsequently performed to assess the validity of statistically significant findings, using slightly different thresholds for MVPA classification, different lengths of Fitbit data, categorical time labels, and different patient inclusion strategies.

GMM was performed using the *lcmm* package [41] (version 2.1.0) in R (version 4.3.3; R Foundation); LME was performed using *lmerTest* [42] (version 3.1-3) in R (version 4.3.3); other analyses were performed using SciPy (version 1.10.1) and statsmodels (version 0.14.0) in Python (version 3.8.15). A P value of less than .05 is considered indicative of a statistically significant difference or association. Due to the exploratory, hypothesis-generating nature of our analyses, we did not make adjustments for multiple testing.

Results

Participant Characteristics

Of the 172 enrolled patients, 80.2% (n=138) were female; 75.6% (n=130) were White; 44.8% (n=77) identified as Hispanic, with 53.5% (n=92) selecting Spanish as their preferred language; 47.7% (n=82) had at least some college education; 45.3% (n=78) were unemployed, and 94.8% (n=163) had an estimated annual income below US $50,000 (Multimedia Appendix 2). Of those enrolled, 82 (47.7%) were classified as valid wear patients, providing 7566 valid wear days in total during the entire study period. Noncompliance with Fitbit wear was related to forgetfulness, rashes, and limited digital proficiency. Table 1 shows the baseline characteristics of valid wear patients whose activity data were analyzed in this work. Compared to those who did not meet the valid wear criteria (n=90, Multimedia Appendix 2), patients with valid wear were younger (mean 46.9, SD 13.0 vs mean 51.4, SD 9.6 years, P=.01), less likely to be Hispanic (37% vs 55%, P=.03) or Spanish-speaking (43% vs 63%, P=.01), and less likely to be married (26% vs 54%, P<.001). No other significant difference was observed in baseline characteristics and PROs between the valid wear and the invalid wear group.

Table 1. Characteristics and COVID experience of valid wear patients at baseline. Valid wear patients had at least 20 hours of valid Fitbit data per day for at least 7 days throughout this study.

Characteristics	Valid wear (n=82)
Age, mean (SD)	46.9 (13.0)
Sex, n (%)	
Female	68 (83)
Male	14 (17)
Race, n (%)	
White	59 (72)
Non-White	13 (16)
Unknown	10 (12)
Ethnicity, n (%)	
Non-Hispanic or Latino	50 (61)
Hispanic	29 (35)
Unknown	3 (4)
Education, n (%)	
Grade 5 or less	6 (7)
Grade 6 to 8	4 (5)
Grade 9 to 12	25 (30)
Some college or college	39 (48)
Postgraduate	5 (6)
Unknown	3 (4)
Employment status, n (%)	
Employed	35 (43)
Unemployed	35 (43)
Student	5 (6)
Retired	2 (2)
Unknown	5 (6)
Housing status, n (%)	
Rent	64 (78)
Own	10 (12)
Homeless	5 (6)
Unknown	3 (4)
Marital status, n (%)	
Single	40 (49)
Married	21 (26)
Other	21 (26)
Estimated annual income (US $), n (%)	
0	25 (30)
1-19,999	30 (37)
20,000-49,999	22 (27)
50,000+	2 (2)
Unknown	3 (4)
Preferred language, n (%)	

Characteristics	Valid wear (n=82)
English	47 (57)
Spanish	35 (43)
Frequency of completing 150 minutes per week of moderate-intensity physical activity before contracting COVID-19, n (%)	
Never	19 (23)
Very few weeks	13 (16)
Some weeks	4 (5)
Most weeks	11 (13)
Every week	28 (34)
Unknown	7 (9)
Frequency of completing 150 minutes per week of vigorous-intensity physical activity before contracting COVID-19, n (%)	
Never	47 (57)
Very few weeks	9 (11)
Some weeks	5 (6)
Most weeks	2 (2)
Every week	13 (16)
Unknown	6 (7)
Admitted to the hospital due to COVID-19, n (%)	17 (21)
Admitted to the ICU[a] due to COVID-19, n (%)	9 (11)
Prescribed supplementary oxygen support due to COVID-19, n (%)	15 (18)
Intubated due to COVID, n (%)	2 (2)
Long COVID and ME/CFS[b] diagnosis, n (%)	
Long COVID only	71 (87)
ME/CFS only	4 (5)
Both long COVID and ME/CFS	6 (7)

[a]ICU: intensive care unit.

[b]ME/CFS: myalgic encephalomyelitis/chronic fatigue syndrome.

Fitbit Findings and Patient Categorization

Fitbit data between baseline and 6-month dates were used for subsequent analyses in consideration of the number of patients contributing valid data and their potential association with PROs at 6-month follow-up. These include 4141 days of Fitbit data from valid wear patients, with each patient providing 50.5 (SD 38.5) days of data on average (median 42, IQR 20-70.75). The distribution of days of valid Fitbit data since Fitbit deployment through the 6-month PRO assessment for each valid wear participant can be found in Multimedia Appendix 3. Noncompliance with PRO completion was associated with survey length, difficulty reaching patients, and symptom-related challenges.

After applying MVPA categorization using average daily MVPA minutes between baseline and 6 months, valid wear patients (n=82) were classified into an MVPA-active subgroup (n=41) and an MVPA-inactive subgroup (n=41). MVPA-inactive patients had an average number of valid wear days of 60.2 (SD 41.9; median 51, IQR 23-87), whereas MVPA-active patients had an average of 40.8 (SD 32.4; median 28, IQR 16-63). The Wilcoxon rank test indicated that the MVPA-inactive group had significantly more valid wear days than the MVPA-active group (P=.33). Table 2 summarizes average daily Fitbit characteristics of the two groups. Consistent with their labels, the MVPA-active group had significantly higher MVPA minutes (P<.001), steps (P<.001), maximum HR (P<.001), and significantly lower percentages of sedentary time (P=.002) than the MVPA-inactive group.

Table 2. Average daily Fitbit characteristics of MVPA[a]-active and MVPA-inactive patients between baseline and 6 months. Patients were categorized by their individual average daily MVPA minutes.

Daily Fitbit characteristics	MVPA-active (n=41), mean (SD)	MVPA-inactive (n=41), mean (SD)	*P* value
Sedentary time/wear time (%)	66.8 (9.8)	74 (8.8)	.002
Lightly active minutes	283.2 (81.0)	242.1 (82.7)	.06
MVPA minutes	46.9 (31.4)	8.2 (6.3)	<.001
Steps	8966 (3818)	5744 (2582)	<.001
Resting heart rate	70.7 (9.0)	70.6 (8.7)	.79
Maximum heart rate	127.8 (8.5)	121.2 (8.1)	<.001
Go-to-bed time	23.99 (1.51)	24.11 (1.70)	.98
Get-up time	7.60 (1.57)	8.01 (1.94)	.37
Hours asleep	6.83 (1.36)	6.96 (1.04)	.91
Asleep time/time in bed (%)	87.4 (2.7)	87.1 (3.7)	.85

[a]MVPA: moderate to vigorous physical activity.

Patient-Reported Outcomes

The 2 groups reported significantly different scores in the physical function, fatigue, pain interference, social roles, and dyspnea scales at baseline (Table 3). The differences in group means all reached the general meaningful change threshold of 3 T-score points for group comparisons [43]. After converting into severity categories (ie, acceptable or mild, moderate concern, or significant concern) using standardized cutoff scores of PROMIS scales, the differences in physical function, fatigue, and dyspnea scores remained statistically significant (*P* values were .03, .006, and .007, respectively), indicating that patients in the MVPA-inactive subgroup experienced more severe problems as reflected by the 3 subscales than those in the MVPA-active subgroup at baseline. Similar differences were observed in the severity scores of related symptoms.

Table 3. PROs[a] of MVPA[b]-active and MVPA-inactive patients at baseline[c].

PROs	MVPA-active (n=41)	MVPA-inactive (n=41)	*P* value
PHQ-2[d,e] (0-6), ≥3[f], n (%)	8 (20)	15 (37)	.08
PROMIS-29[g] depression[e] (20-80), mean (SD)	53 (10)	56 (11)	.21
GAD-7[e,h] (0-21), mean (SD)	6.1 (5.0)	8.2 (5.7)	.12
PROMIS-29 anxiety[e] (20-80), mean (SD)	56 (11)	57 (10)	.77
PROMIS-29 physical function[i] (20-80), mean (SD)	38 (7)	34 (7)	.006
PROMIS-29 fatigue[e] (20-80), mean (SD)	59 (10)	66 (8)	<.001
PROMIS-29 sleep disturbance[e] (20-80), mean (SD)	58 (5)	58 (6)	.55
PROMIS-29 pain interference[e] (20-80), mean (SD)	60 (10)	65 (9)	.03
PROMIS-29 social roles[i] (20-80), mean (SD)	47 (11)	40 (10)	.02
PROMIS[j] dyspnea[e] (20-80), mean (SD)	56 (10)	63 (12)	.002
PROMIS cognition[i] (20-80), mean (SD)	46 (12)	41 (10)	.06
Fatigue, tiredness, or exhaustion			
Yes, n (%)	36 (88)	39 (95)	.43
Severity score[k] (0-16), mean (SD)	9.2 (4.2)	12.3 (4.2)	.001
Muscle pain, muscle cramps, or muscle aches			
Yes, n (%)	33 (80)	32 (78)	.99
Severity score[k] (0-16), mean (SD)	7.0 (4.9)	8.1 (5.7)	.37
Joint pain			
Yes, n (%)	25 (61)	29 (71)	.48
Severity score[k] (0-16), mean (SD)	5.6 (5.3)	7.5 (5.5)	.15
Unrefreshing sleep			
Yes, n (%)	29 (71)	34 (83)	.30
Severity score[k] (0-16), mean (SD)	7.1 (6.0)	9.8 (5.3)	.06
Problems getting to sleep, sleeping through the night, or waking up on time			
Yes, n (%)	34 (83)	37 (90)	.52
Severity score[k] (0-16), mean (SD)	7.1 (5.7)	10.4 (4.7)	.007
Forgetfulness or memory problems			
Yes, n (%)	24 (59)	31 (76)	.16
Severity score[k] (0-16), mean (SD)	5.2 (5.6)	8.1 (5.5)	.03
Difficulty thinking or concentrating			
Yes, n (%)	26 (63)	35 (85)	.04
Severity score[k] (0-16), mean (SD)	5.3 (5.2)	8.7 (5.4)	.006
Dizziness or lightheadedness			
Yes, n (%)	29 (71)	33 (80)	.44
Severity score[k] (0-16), mean (SD)	4.8 (4.5)	4.4 (3.6)	.98
Post-exertional malaise[l], yes, n (%)	32 (78)	30 (73)	.80

[a]PRO: patient-reported outcome.

[b]MVPA: moderate to vigorous physical activity.

[c]Ranges of patient-reported outcome scores are indicated in parentheses. PROMIS T-scores have a mean of 50 and an SD of 10 in the reference population and typically fall between 20 and 80.

[d]PHQ-2: Patient Health Questionnaire-2.

[e]On PHQ-2, GAD-7, and all PROMIS subscales except physical function, social roles, and cognition, a higher score indicates more severe problems or concerns.

[f]A PHQ-2 score of 3 or greater indicates that major depressive disorder is likely.

[g]PROMIS-29: Patient-Reported Outcomes Measurement Information System-29.

[h]GAD-7: Generalized Anxiety Disorder-7.

[i]On PROMIS physical function, social roles, and cognition subscales, a lower score indicates more severe problems or concerns.

[j]PROMIS: Patient-Reported Outcomes Measurement Information System.

[k]Symptom severity scores were calculated using the scoring algorithm of the CDC Symptom Inventory [34,35]. A higher score indicates that the symptom is more frequent or more intense. A score of 0 means no symptom.

[l]Post-exertional malaise was measured by a yes, no, or do not know question asking whether any of the symptoms get worse for at least 24 hours after activity.

Table 4 provides average individual changes in PROs from baseline to 3 months and 6 months, showing the direction and magnitude of the trends in PROs. However, the effect of other factors is not considered, and the sample size is limited due to the pairwise within-subject calculation. To address this, we further analyzed longitudinal associations between MVPA levels and PROs with LME regression, which is able to handle missing values while modeling the effects of multiple factors.

Table 4. Average individual changes in PROs[a] from baseline to 3 and 6 months[b].

PRO change scores	From baseline to 3 months		From baseline to 6 months	
	MVPA[c]-active	MVPA-inactive	MVPA-active	MVPA-inactive
PHQ-2[d,e] (0-6), mean (SD)	+0.2 (1.7)	–0.3 (1.8)	+0.2 (1.4)	–0.8 (2.0)
PROMIS-29[f] depression[e] (20-80), mean (SD)	+1 (9)	0 (9)	+2 (8)	–2 (9)
GAD-7[e,g] (0-21), mean (SD)	+0.8 (5.2)	–0.7 (4.6)	+1.4 (4.6)	–1.6 (6.0)
PROMIS-29 anxiety[e] (20-80), mean (SD)	–4 (11)	–1 (12)	–1 (6)	–2 (11)
PROMIS-29 physical function[h] (20-80), mean (SD)	–1 (6)	+1 (6)	–1 (8)	+2 (7)
PROMIS-29 fatigue[e] (20-80), mean (SD)	0 (9)	–2 (9)	+2 (10)	–3 (9)
PROMIS-29 sleep disturbance[e] (20-80), mean (SD)	–1 (6)	–1 (7)	0 (8)	–2 (7)
PROMIS-29 pain interference[e] (20-80), mean (SD)	+1 (10)	–1 (11)	0 (10)	–2 (9)
PROMIS-29 social roles[h] (20-80), mean (SD)	0 (11)	+3 (10)	0 (11)	+8 (10)
PROMIS dyspnea[e] (20-80), mean (SD)	+1 (12)	–4 (13)	–1 (10)	–2 (13)
PROMIS cognition[h] (20-80), mean (SD)	–1 (9)	+2 (9)	0 (13)	+4 (9)

[a]PRO: patient-reported outcome.

[b]Ranges of PRO scores are indicated in parentheses. PROMIS T-scores have a mean of 50 and an SD of 10 in the reference population and typically fall between 20 and 80. Data from symptom severity scores is not included due to limited sample sizes, since only those who have reported a symptom would receive follow-up questions about that symptom in the next round of questionnaires.

[c]MVPA: moderate to vigorous physical activity.

[d]PHQ-2: Patient Health Questionnaire-2.

[e]On PHQ-2, GAD-7, and all PROMIS subscales except physical function, social roles, and cognition, a higher score indicates more severe problems or concerns.

[f]PROMIS-29: Patient-Reported Outcomes Measurement Information System-29.

[g]GAD-7: Generalized Anxiety Disorder-7.

[h]On PROMIS physical function, social roles, and cognition subscales, a lower score indicates more severe problems or concerns.

Longitudinal Associations Between MVPA and PROs

LME analysis indicated a statistically significant effect of MVPA level on scores of several PROs over time. For every 3 months of being MVPA-active compared with MVPA-inactive, the coefficient estimates are –1.94 (95% CI –3.05 to –0.82, *P*<.001) for raw scores of PROMIS-29 ability to participate in social roles and activities scale (abbreviated as PROMIS-29 social roles) and –4.21 (95% CI –6.64 to –1.78, *P*<.001) for its T-scores. The latter is greater than both the general meaningful change threshold of 3 T-score points for group comparison on PROMIS scales and the minimal important change values of 0.4-2.2 T-score points estimated by previous studies for the PROMIS-29 social roles scale [43,44]. For raw physical function scores, this coefficient is –1.00 (95% CI –1.97 to –0.04, *P*=.04) but is no longer significant when regressing on T-scores (estimate=–1.35, 95% CI –3.00 to 0.30, *P*=.11). Considering that a lower score indicates more severe problems on both physical function and social roles subscales, these negative coefficient estimates suggest a potential negative effect of high MVPA levels or a beneficial effect of low MVPA levels over time. Aligning with LME regression results, average individual change scores (Table 4) on physical function and social roles scales show that the scores of MVPA-active patients remained around the same on average, while patients in the MVPA-inactive group had increased scores on average. Together, these findings indicate that MVPA-inactive patients experienced more improvements in outcomes on average after 6 months compared to MVPA-active patients, especially in the ability to participate in social roles.

LME regression of long COVID symptom severity scores identified a statistically significant effect of MVPA group over time on symptoms related to sleep quality (ie, problems getting to sleep, sleeping through the night, or waking up on time), with a coefficient estimate of 2.06 (95% CI 0.40 to 3.71, *P*=.02) for being MVPA-active over 3 months. Similar to the observations from PROMIS-29 scores, this positive estimate suggests that MVPA-inactive patients had more improvements in sleep quality after 6 months than MVPA-active patients. However, this effect was not reflected by the PROMIS-29 sleep disturbance scale.

The effect of MVPA level over time is not significant for any other PROs. In addition, the main effects of MVPA groups at baseline in LME regression are consistent with baseline comparisons in Table 3 when baseline PRO scores are not included as a predictor. The addition of other Fitbit or demographic variables did not result in improved model fitness. The only new variable that appeared to be statistically significant was step count, which is highly correlated with MVPA minutes. Model diagnostic tests indicated that the fitted models were valid. Subsequent sensitivity analysis using slightly different thresholds for the classification of MVPA groups and data inclusion in LME regression produced the same findings, with only small fluctuations in coefficient estimates and *P* values. For example, 2 patients had average daily MVPA levels near the classification cutoff. After moving them to the other group, the coefficient estimate of MVPA-time interaction remained significant. Removal of patients with near-cutoff numbers of valid wear days and overly skewed or scattered valid wear days (eg, patient 039, 064, and 137) produced the same findings. We also tried removing 9 patients who had valid PRO scores at only one time point, and findings remained the same. PRO data from patients who missed 1 time point were considered valid and included in LME models.

Exploratory Analyses

The GMM of day-level Fitbit time series identified 1 cluster of resting HR, 1 cluster of step counts, and 2 clusters of MVPA minutes. The GMM model of daily resting HR has an intercept of 71.5 (*P*<.001) beats per minute, and a slope of –0.023 (*P*=.01) beats per minute per day, indicating a slight decreasing trend in resting HR over time on average. The intercept represents the modeled average value on day 0 of Fitbit deployment under specified model settings and model constraints. The step count model has an intercept of 7289 steps (*P*<.001) and a slope of 2.51 (*P*=.53) steps per day, indicating that the average step count of the patients was stable over time in general. The 2 clusters of daily MVPA minutes have sizes of 70 and 12 patients, intercepts of 18.4 and 68.2 minutes (both with *P*<.001), and slopes of –0.029 (*P*=.46) and 0.20 (*P*=.39) minutes per day, respectively. This indicates that there were an MVPA-active subgroup and an MVPA-inactive subgroup, but the activities of both subgroups did not exhibit a clear trend over time. As GMM does not provide specific cutoffs for patient clusters and GMM-identified clusters may be of insufficient sample size, we did not use GMM-identified clusters for further analysis, but instead selected the WHO guideline, which is more reusable and comparable for future studies. Due to concerns over the missingness and quality of the sleep data from Fitbit, we did not analyze this data.

Further analyses were conducted to study potential factors related to the observed MVPA levels. Employment status at baseline was similar between MVPA-active and MVPA-inactive groups, and differences in Fitbit measures remained significant after adjusting for employment in a 2-way ANOVA, indicating that MVPA differences were not explained by employment status alone. Multivariate linear regression of MVPA minutes against demographic characteristics and baseline survey scores, with either a logistic model or a 0-inflated negative binomial model, revealed sex and baseline PROMIS-29 fatigue score as significant predictors. Specifically, males and individuals with less severe fatigue at baseline were correlated with being MVPA-active during this study's period.

Other exploratory analyses on the Fitbit data, including dimensionality reduction of average Fitbit features (to study potential patient clusters), time series clustering of minute-level data (to study daily activity patterns), weekday-to-weekend and weekly comparisons (to study weekly patterns), and alignment of patient activities on adjacent days (to study day-to-day variations or potential activity pacing behaviors), did not produce meaningful findings beyond reported results and existing literature.

Discussion

Principal Findings

This study collected Fitbit-measured physical activity data and PROs from a subset of long COVID and ME/CFS patients in

the LC&FIRP study to identify longitudinal patterns and examine the overlay of PROs in relation to patient characteristics. The results showed that patients could remain MVPA-active despite experiencing symptoms. According to the Physical Activity Guidelines for Americans, moderate housework or yardwork, such as carrying groceries or raking the yard, may reach the threshold of moderate activity [45]. One of the main characteristics of long COVID is the impact of symptoms that can alter functionality within these critical everyday tasks [46]. While we do not have records of why or how some patients maintained high activity levels, one factor to consider is the socioeconomic status of our patient cohort. FHCSD is one of the ten largest FQHCs in the nation, and the vast majority of patients are low-income [47]. It is important to analyze data across diverse patient populations, as those at an FQHC may have experienced household or financial demands, including the need to continue working. Further study of employment type and partner's employment may provide additional insight. These factors highlight both the need for pacing, planning, and prioritizing as well as the real challenges that prevent their implementation [48].

In long COVID, where fatigue and PEM are common, physical activity should be recommended with care in an individualized approach. In our study, 75.6% (62/82) of patients reported PEM, and the proportions in the 2 MVPA subgroups were not significantly different (P=.80). The high rate of PEM may reflect our patient population's responsibilities when it comes to providing and caring for their families. Although some studies have demonstrated improvements in metrics related to physical function after physical rehabilitation programs, few addressed PEM, and those that did found no improvement in exercise performance following structured exercise interventions [49-54].

In our attempts to quantify PEM, we found that it was difficult to establish parameters due to the diversity in patient experiences. Although symptoms of PEM vary, fatigue, cognitive dysfunction, and sleep problems are reported with high frequency [55]. Velez-Santamaria et al [46] examined the changes in fatigue severity in long COVID participants, finding that it worsened significantly, with 92.4% of participants meeting the criteria for ME/CFS, while Stussman et al [56] highlighted the diverse onset, duration, and recovery patterns of PEM symptoms. Similar to our study, this heterogeneity underscores the need to tailor treatment strategies based on the severity of fatigue, symptoms, and the patient's lifestyle [56].

Of the 11 published studies on long COVID and wearable devices, ours was the only one to use PROMIS-29 as the primary PRO measure [26]. We found that at baseline, MVPA-inactive patients experienced more severe problems with physical function, fatigue, dyspnea, and sleep quality than MVPA-active patients. On average, after 6 months, MVPA-inactive patients reported some recovery in the domains of sleep quality and the ability to participate in social roles, compared to MVPA-active patients whose metrics worsened or remained around the same. Other studies used different PROS, including the Fatigue Assessment Scale [49], Short Form 36 Health Survey [50], or the COVID-19 Yorkshire Rehabilitation Scale [57], making comparisons challenging due to heterogeneity in scales, symptom emphasis, and reference periods. As PROMIS-29 was not used in prior studies, direct trajectory comparisons with other cohorts are limited.

Previous studies also examined the association between physical activity and patient with long COVID's experiences. Humphreys et al [58] conducted qualitative interviews with patients with long COVID, with one of the main themes being the struggle with impaired physical function. In a cross-sectional study by Wright et al [59], about 75% of the participants reported that physical activity worsened their long COVID symptoms. A later cross-sectional study by Vélez-Santamaría et al [46], involving multiple quantitative questionnaires similar to our study, revealed a significant association between impaired functionality, lower physical activity levels, and worsened quality of life in patients with long COVID. Rekeland et al [60] also reported similar associations between physical activity, ME/CFS severity, and PROs in patients with ME/CFS. Findings in our study are consistent with and complementary to those studies. To our knowledge, our study is the first to report longitudinal patterns of physical activity in lower-income patients with long COVID using a wearable device and PROs. Other studies exist, but do not focus on similar cohorts [61-63].

Our finding that a lower MVPA level was associated with more improvements in outcomes in the long-term seems to suggest some benefit from reduced activity. Given our observational results, we propose that it may be beneficial for health professionals to ask patients about their normal pattern of physical activity and symptoms that follow or worsen after activity. If the patient's responsibilities require a substantial amount of MVPA, strategies for mitigating its potentially negative effects on symptoms could be explored. One potential strategy is activity pacing, that is, dividing physical activities into multiple portions that are more manageable and balancing them with rests [64]. In our study, we did not observe clear indications in the exploratory analysis (result not shown) that patients paced their activities on the day level because they maintained a relatively stable amount of MVPA on weekdays over time. In addition, our PROs were only conducted every 3 months, which may have resulted in difficulty for the patient to recall certain activities that caused PEM. Due to the exploratory nature of our study, results should be interpreted as hypothesis-generating.

The findings of this study should be interpreted in light of its limitations. First, the sample size of this study was relatively small (n=82 for valid wear patients) and was observational in nature, which limits the extent to which results can be generalized. Second, exploratory analysis is typically used to generate hypotheses of effects and associations rather than to confirm them definitively. The goal is often to identify potential patterns, associations, or areas for further study, rather than to make conclusions about the data. It is important to interpret the P values reported with caution, given that the number of statistical tests performed increases the risk of type I errors (false positives). As such, our findings provide insights into the longitudinal trajectory of patients with long COVID and should be used to guide future research. Third, due to concerns over the missingness and quality of Fitbit data on sleep and HRV, we were unable to investigate these characteristics. Future studies could consider collecting these measures through other

wearable devices or even including biological markers to better understand these behavioral and physiological aspects among patients with long COVID and related IACCs. Fourth, as a substudy of a large intervention program, this study did not contain a healthy control group or other illness comparison groups, unlike other studies [62,63]. Therefore, patients were compared to themselves in the past or to other patients with different characteristics. Fifth, the applicability of the WHO MVPA guideline to patients with long COVID's needs further validation. Lastly, our sample was focused on patients at a single site, which may affect generalizability to broader populations, whereas other studies had a more diverse population [26].

Considering that repeated measures studies could be undermined by the regression to the mean phenomenon, we conducted baseline-adjusted LME regression analysis, which appropriately accounts for patient-level variability in repeated measures. Moreover, the major effect we observed on the PROMIS-29 social roles scale is greater than the meaningful change thresholds of the PROMIS scales, which is unlikely to be solely due to regression to the mean. Nevertheless, a large, real-world interventional cohort that includes a diverse set of patients may help phenotype patients in a more comprehensive manner, potentially leading to a better understanding of the disease, treatment strategies, and the development of effective interventions. Lastly, our project lacked the ability to monitor and follow up with patients in real time when they were experiencing a worsening symptom or PEM. The use of ecological momentary assessment is a crucial next step to better comprehend the sporadic nature of symptoms and PEM in patients with the goal of gathering information about daily patterns of symptoms to identify triggers for the worsening of symptoms [65].

Conclusions

In conclusion, our results highlight the complex nature and diverse impact of long COVID and ME/CFS, drawing attention to the value of combining self-reported symptoms and objective physical activity data when evaluating individuals. Furthermore, the results suggest that disease experience is individualized. As this study was exploratory in nature, our findings could help lay the groundwork for clinician-patient interactions and for tailoring rehabilitation efforts and guiding future research. Our paper highlights the importance of continuing research in the field of long COVID and ME/CFS. Next steps could include a deeper dive into physiological aspects such as resting metabolic rate, ecological momentary assessment, HRV, and sleep over a continuous period. Using a combination of different wearables in a larger sample size could help to further the understanding of personal trajectories of long COVID experience, hopefully leading to personalized treatment options and better tracking of disease progression.

Acknowledgments

Foremost, the authors would like to thank the Long COVID and Fatiguing Illness Recovery patients for their time and willingness to be a part of this study. The authors are grateful to the numerous staff members and administrative leadership at Family Health Centers of San Diego for providing in-kind support for the development of this study.

Funding

This study was made possible by funding from the Conrad Prebys Foundation and the Centers for Disease Control and Prevention (75D30121C11341). The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the CDC.

Data Availability

The deidentified original datasets used in this study are available from the corresponding author upon reasonable request. Due to ethical, legal, and privacy considerations, we refrain from sharing the original data through a public repository. Necessary summary data for study findings have been provided in the main text or as supplementary materials.

Conflicts of Interest

None declared.

Multimedia Appendix 1
STROBE checklist for cohort studies.
[DOCX File , 41 KB - formative_v9i1e77644_app1.docx]

Multimedia Appendix 2
Complete patient characteristics and PROs at baseline. PRO: patient reported outcomes.
[DOCX File , 58 KB - formative_v9i1e77644_app2.docx]

Multimedia Appendix 3

Distributions of valid wear days of each valid wear patient between Fitbit deployment and 6-month PRO assessment. Each point marks a valid wear day. Patients are colored by their MVPA group. The x-axis does not correspond to survey dates. MVPA: moderate to vigorous physical activity; PRO: patient reported outcomes.
[PNG File , 660 KB - formative_v9i1e77644_app3.png]

References

1. COVID-19 deaths. WHO COVID-19 Dashboard. URL: https://data.who.int/dashboards/covid19/deaths [accessed 2025-12-06]
2. COVID-19 cases. WHO COVID-19 dashboard. URL: https://data.who.int/dashboards/covid19/cases [accessed 2025-12-06]
3. Williams S, Sheard N, Stuart B, Phan HT, Borca F, Wilkinson TM, REACT COVID Investigators. Comparisons of early and late presentation to hospital in COVID-19 patients. Respirology 2021;26(2):204-205 [FREE Full text] [doi: 10.1111/resp.13985] [Medline: 33283433]
4. Nalbandian A, Sehgal K, Gupta A, Madhavan MV, McGroder C, Stevens JS, et al. Post-acute COVID-19 syndrome. Nat Med 2021 Apr;27(4):601-615 [FREE Full text] [doi: 10.1038/s41591-021-01283-z] [Medline: 33753937]
5. Huerne K, Filion KB, Grad R, Ernst P, Gershon AS, Eisenberg MJ. Epidemiological and clinical perspectives of long COVID syndrome. Am J Med Open 2023;9:100033 [FREE Full text] [doi: 10.1016/j.ajmo.2023.100033] [Medline: 36685609]
6. National Academies of Sciences, Engineering, and Medicine. In: Goldowitz I, Worku T, Brown L, V. Fineberg H, editors. A Long COVID Definition: A Chronic, Systemic Disease State With Profound Consequences. Washington (DC): National Academies Press; 2024. [doi: 10.17226/27768]
7. Ford ND, Agedew A, Dalton AF, Singleton J, Perrine CG, Saydah S. Notes from the field: long COVID prevalence among adults - United States, 2022. MMWR Morb Mortal Wkly Rep 2024;73(6):135-136 [FREE Full text] [doi: 10.15585/mmwr.mm7306a4] [Medline: 38359012]
8. Long COVID - household pulse survey - COVID-19. 2025. URL: https://www.cdc.gov/nchs/covid19/pulse/long-covid.htm [accessed 2025-12-06]
9. Bayarri-Olmos R, Bain W, Iwasaki A. The role of complement in long COVID pathogenesis. JCI Insight 2025;10(16):e194314 [FREE Full text] [doi: 10.1172/jci.insight.194314] [Medline: 40857408]
10. National Academies of Sciences, Engineering, and Medicine. Toward a Common Research Agenda in Infection-Associated Chronic Illnesses: Proceedings of a Workshop. Washington (DC): National Academies Press (US); 2024. [doi: 10.17226/27462]
11. Komaroff AL, Lipkin WI. ME/CFS and long COVID share similar symptoms and biological abnormalities: road map to the literature. Front Med (Lausanne) 2023;10:1187163 [FREE Full text] [doi: 10.3389/fmed.2023.1187163] [Medline: 37342500]
12. Shabnam S, Razieh C, Dambha-Miller H, Yates T, Gillies C, Chudasama YV, et al. Socioeconomic inequalities of long COVID: a retrospective population-based cohort study in the United Kingdom. J R Soc Med 2023;116(8):263-273 [FREE Full text] [doi: 10.1177/01410768231168377] [Medline: 37164035]
13. Subramanian A, Nirantharakumar K, Hughes S, Myles P, Williams T, Gokhale KM, et al. Symptoms and risk factors for long COVID in non-hospitalized adults. Nat Med 2022;28(8):1706-1714 [FREE Full text] [doi: 10.1038/s41591-022-01909-w] [Medline: 35879616]
14. Lukkahatai N, Rodney T, Ling C, Daniel B, Han H. Long COVID in the context of social determinants of health. Front Public Health 2023;11:1098443 [FREE Full text] [doi: 10.3389/fpubh.2023.1098443] [Medline: 37056649]
15. Xiang J, Zheng H, Cai Y, Chen S, Wang Y, Chen R. Cumulative social disadvantage and its impact on long COVID: insights from a U.S. national survey. BMC Med 2025;23(1):207 [FREE Full text] [doi: 10.1186/s12916-025-04039-5] [Medline: 40189508]
16. Davis HE, Assaf GS, McCorkell L, Wei H, Low RJ, Re'em Y, et al. Characterizing long COVID in an international cohort: 7 months of symptoms and their impact. eClinicalMedicine 2021;38:101019 [FREE Full text] [doi: 10.1016/j.eclinm.2021.101019] [Medline: 34308300]
17. Burton C, Dawes H, Goodwill S, Thelwell M, Dalton C. Within and between-day variation and associations of symptoms in long COVID: intensive longitudinal study. PLoS One 2023;18(1):e0280343 [FREE Full text] [doi: 10.1371/journal.pone.0280343] [Medline: 36656830]
18. Asahina K, de Bivort BL, Grunwald Kadow IC, Yapici N. Editorial: revisiting behavioral variability: what it reveals about neural circuit structure and function. Front Behav Neurosci 2022;16:956388 [FREE Full text] [doi: 10.3389/fnbeh.2022.956388] [Medline: 35783229]
19. Dhingra LS, Aminorroaya A, Oikonomou EK, Nargesi AA, Wilson FP, Krumholz HM, et al. Use of wearable devices in individuals with or at risk for cardiovascular disease in the US, 2019 to 2020. JAMA Netw Open 2023;6(6):e2316634 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.16634] [Medline: 37285157]
20. Mishra T, Wang M, Metwally AA, Bogu GK, Brooks AW, Bahmani A, et al. Pre-symptomatic detection of COVID-19 from smartwatch data. Nat Biomed Eng 2020;4(12):1208-1220 [FREE Full text] [doi: 10.1038/s41551-020-00640-6] [Medline: 33208926]

21. Behar JA, Liu C, Kotzen K, Tsutsui K, Corino VDA, Singh J, et al. Remote health diagnosis and monitoring in the time of COVID-19. Physiol Meas 2020;41(10):10TR01 [FREE Full text] [doi: 10.1088/1361-6579/abba0a] [Medline: 32947271]
22. Jeong H, Rogers JA, Xu S. Continuous on-body sensing for the COVID-19 pandemic: gaps and opportunities. Sci Adv 2020;6(36):eabd4794 [FREE Full text] [doi: 10.1126/sciadv.abd4794] [Medline: 32917604]
23. Khondakar KR, Kaushik A. Role of wearable sensing technology to manage long COVID. Biosensors (Basel) 2022;13(1):62 [FREE Full text] [doi: 10.3390/bios13010062] [Medline: 36671900]
24. Hirten RP, Danieletto M, Tomalin L, Choi KH, Zweig M, Golden E, et al. Use of physiological data from a wearable device to identify SARS-CoV-2 infection and symptoms and predict COVID-19 diagnosis: observational study. J Med Internet Res 2021;23(2):e26107 [FREE Full text] [doi: 10.2196/26107] [Medline: 33529156]
25. Singh B, Chastin S, Miatke A, Curtis R, Dumuid D, Brinsley J, et al. Real-world accuracy of wearable activity trackers for detecting medical conditions: systematic review and meta-analysis. JMIR mHealth uHealth 2024;12:e56972 [FREE Full text] [doi: 10.2196/56972] [Medline: 39213525]
26. Kukreti S, Lu M, Yeh C, Ko N. Physiological sensors equipped in wearable devices for management of long COVID persisting symptoms: scoping review. J Med Internet Res 2025;27:e69506 [FREE Full text] [doi: 10.2196/69506] [Medline: 40137051]
27. Goosen A, Foster-Bonds R, Vogel JM. Wearable devices enable long COVID patients to decrease symptom severity: a case series from pilot user testing. Cardiopulm Phys Ther J 2025;36(2):99-104. [doi: 10.1097/CPT.0000000000000268] [Medline: 40190996]
28. Godino JG, Samaniego JC, Sharp SP, Taren D, Zuber A, Armistad AJ, et al. A technology-enabled multi-disciplinary team-based care model for the management of long COVID and other fatiguing illnesses within a federally qualified health center: protocol for a two-arm, single-blind, pragmatic, quality improvement professional cluster randomized controlled trial. Trials 2023;24(1):524 [FREE Full text] [doi: 10.1186/s13063-023-07550-3] [Medline: 37573421]
29. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med 2007;147(8):573-577 [FREE Full text] [doi: 10.7326/0003-4819-147-8-200710160-00010] [Medline: 17938396]
30. Fieo R, Ocepek-Welikson K, Kleinman M, Eimicke JP, Crane PK, Cella D, et al. Measurement equivalence of the Patient Reported Outcomes Measurement Information System (PROMIS) applied cognition - general concerns, short forms in ethnically diverse groups. Psychol Test Assess Model 2016;58(2):255-307 [FREE Full text] [Medline: 28523238]
31. PROMIS SF v1.0-Dyspnea-Severity 10a. FACIT Group. URL: https://www.facit.org/measures/promis-sf-v1.0-dyspnea-severity-10a [accessed 2025-12-06]
32. Patient health questionnaire (PHQ) screeners. URL: https://www.phqscreeners.com/select-screener/ [accessed 2025-12-06]
33. Huang W, Rose AJ, Bayliss E, Baseman L, Butcher E, Garcia R, et al. Adapting summary scores for the PROMIS-29 v2.0 for use among older adults with multiple chronic conditions. Qual Life Res 2019;28(1):199-210 [FREE Full text] [doi: 10.1007/s11136-018-1988-z] [Medline: 30209722]
34. Wagner D, Nisenbaum R, Heim C, Jones JF, Unger ER, Reeves WC. Psychometric properties of the CDC Symptom Inventory for assessment of chronic fatigue syndrome. Popul Health Metr 2005;3:8 [FREE Full text] [doi: 10.1186/1478-7954-3-8] [Medline: 16042777]
35. Unger ER, Lin JS, Wisk LE, Yu H, L'Hommedieu M, Lavretsky H, Innovative Support for Patients with SARS-CoV-2 Infections Registry (INSPIRE) Group. Myalgic encephalomyelitis/chronic fatigue syndrome after SARS-CoV-2 Infection. JAMA Netw Open 2024;7(7):e2423555 [FREE Full text] [doi: 10.1001/jamanetworkopen.2024.23555] [Medline: 39046739]
36. Wing D, Godino JG, Baker FC, Yang R, Chevance G, Thompson WK, et al. Recommendations for identifying valid wear for consumer-level wrist-worn activity trackers and acceptability of extended device deployment in children. Sensors (Basel) 2022;22(23):9189 [FREE Full text] [doi: 10.3390/s22239189] [Medline: 36501894]
37. HealthMeasures scoring service. Assessment Center. URL: https://www.assessmentcenter.net/ac_scoringservice [accessed 2025-12-06]
38. Bull FC, Al-Ansari SS, Biddle S, Borodulin K, Buman MP, Cardon G, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med 2020;54(24):1451-1462 [FREE Full text] [doi: 10.1136/bjsports-2020-102955] [Medline: 33239350]
39. Ram N, Grimm KJ. Growth mixture modeling: a method for identifying differences in longitudinal change among unobserved groups. Int J Behav Dev 2009;33(6):565-576 [FREE Full text] [doi: 10.1177/0165025409343765] [Medline: 23885133]
40. Gałecki A, Burzykowski T. Linear Mixed-Effects Model. In: Linear Mixed-Effects Models Using R. New York, NY: Springer; 2013:245-273. [doi: 10.1007/978-1-4614-3900-4_13]
41. Proust-Lima C, Philipps V, Liquet B. Estimation of extended mixed models using latent classes and latent processes: the R Package lcmm. J Stat Soft 2017;78(2):1-56. [doi: 10.18637/jss.v078.i02]
42. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest Package: tests in linear mixed effects models. J Stat Soft 2017;82(13):1-26. [doi: 10.18637/jss.v082.i13]
43. Meaningful Change for PROMIS. HealthMeasures. 2024. URL: https://www.healthmeasures.net/score-and-interpret/interpret-scores/promis/meaningful-change [accessed 2025-12-06]

44. Terwee CB, Peipert JD, Chapman R, Lai J, Terluin B, Cella D, et al. Minimal important change (MIC): a conceptual clarification and systematic review of MIC estimates of PROMIS measures. Qual Life Res 2021;30(10):2729-2754 [FREE Full text] [doi: 10.1007/s11136-021-02925-y] [Medline: 34247326]
45. Piercy KL, Troiano RP, Ballard RM, Carlson SA, Fulton JE, Galuska DA, et al. The physical activity guidelines for Americans. JAMA 2018;320(19):2020-2028 [FREE Full text] [doi: 10.1001/jama.2018.14854] [Medline: 30418471]
46. Vélez-Santamaría R, Fernández-Solana J, Méndez-López F, Domínguez-García M, González-Bernal JJ, Magallón-Botaya R, et al. Functionality, physical activity, fatigue and quality of life in patients with acute COVID-19 and long COVID infection. Sci Rep 2023;13(1):19907 [FREE Full text] [doi: 10.1038/s41598-023-47218-1] [Medline: 37963962]
47. Impact of the health center program. Health Resources & Services Administration. URL: https://bphc.hrsa.gov/about-health-center-program/impact-health-center-program [accessed 2025-12-06]
48. How to conserve your energy. Shared Health. URL: https://sharedhealthmb.ca/files/post-covid-conserving-energy.pdf [accessed 2025-12-06]
49. Kerling A, Beyer S, Dirks M, Scharbau M, Hennemann A, Dopfer-Jablonka A, et al. Effects of a randomized-controlled and online-supported physical activity intervention on exercise capacity, fatigue and health related quality of life in patients with post-COVID-19 syndrome. BMC Sports Sci Med Rehabil 2024;16(1):33 [FREE Full text] [doi: 10.1186/s13102-024-00817-5] [Medline: 38308307]
50. Laguarta-Val S, Varillas-Delgado D, Lizcano-Álvarez Á, Molero-Sánchez A, Melian-Ortiz A, Cano-de-la-Cuerda R, et al. Effects of Aerobic exercise therapy through nordic walking program in lactate concentrations, fatigue and quality-of-life in patients with long-COVID syndrome: a non-randomized parallel controlled trial. J Clin Med 2024;13(4):1035 [FREE Full text] [doi: 10.3390/jcm13041035] [Medline: 38398348]
51. McGregor G, Sandhu H, Bruce J, Sheehan B, McWilliams D, Yeung J, et al. Clinical effectiveness of an online supervised group physical and mental health rehabilitation programme for adults with post-COVID-19 condition (REGAIN study): multicentre randomised controlled trial. BMJ 2024;384:e076506 [FREE Full text] [doi: 10.1136/bmj-2023-076506] [Medline: 38325873]
52. Colas C, Le Berre Y, Fanget M, Savall A, Killian M, Goujon I, et al. Physical activity in long COVID: a comparative study of exercise rehabilitation benefits in patients with long COVID, coronary artery disease and fibromyalgia. Int J Environ Res Public Health 2023;20(15):6513 [FREE Full text] [doi: 10.3390/ijerph20156513] [Medline: 37569053]
53. Philip KEJ, Owles H, McVey S, Pagnuco T, Bruce K, Brunjes H, et al. An online breathing and wellbeing programme (ENO Breathe) for people with persistent symptoms following COVID-19: a parallel-group, single-blind, randomised controlled trial. Lancet Respir Med 2022;10(9):851-862 [FREE Full text] [doi: 10.1016/S2213-2600(22)00125-4] [Medline: 35489367]
54. Pouliopoulou DV, Macdermid JC, Saunders E, Peters S, Brunton L, Miller E, et al. Rehabilitation interventions for physical capacity and quality of life in adults with post-COVID-19 condition: a systematic review and meta-analysis. JAMA Netw Open 2023;6(9):e2333838 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.33838] [Medline: 37725376]
55. Mateo LJ, Chu L, Stevens S, Stevens J, Snell CR, Davenport T, et al. Post-exertional symptoms distinguish myalgic encephalomyelitis/chronic fatigue syndrome subjects from healthy controls. WORK: J Prev, Assess Rehabil 2020;66(2):265-275. [doi: 10.3233/WOR-203168] [Medline: 32568143]
56. Stussman B, Williams A, Snow J, Gavin A, Scott R, Nath A, et al. Characterization of post-exertional malaise in patients with myalgic encephalomyelitis/chronic fatigue syndrome. Front Neurol 2020;11:1025 [FREE Full text] [doi: 10.3389/fneur.2020.01025] [Medline: 33071931]
57. Corrado J, Iftekhar N, Halpin S, Li M, Tarrant R, Grimaldi J, et al. HEART rate variability biofeedback for long COVID Dysautonomia (HEARTLOC): results of a feasibility study. Adv Rehabil Sci Pract 2024;13:27536351241227261 [FREE Full text] [doi: 10.1177/27536351241227261] [Medline: 38298551]
58. Humphreys H, Kilby L, Kudiersky N, Copeland R. Long COVID and the role of physical activity: a qualitative study. BMJ Open 2021;11(3):e047632 [FREE Full text] [doi: 10.1136/bmjopen-2020-047632] [Medline: 33692189]
59. Wright J, Astill SL, Sivan M. The relationship between physical activity and long COVID: a cross-sectional study. Int J Environ Res Public Health 2022;19(9):5093 [FREE Full text] [doi: 10.3390/ijerph19095093] [Medline: 35564488]
60. Rekeland IG, Sørland K, Bruland O, Risa K, Alme K, Dahl O, Fluge. Activity monitoring and patient-reported outcome measures in myalgic encephalomyelitis/chronic fatigue syndrome patients. PLoS One 2022;17(9):e0274472 [FREE Full text] [doi: 10.1371/journal.pone.0274472] [Medline: 36121803]
61. Zahouani I, Desmeules F, Perreault K, Campeau-Lecours A, Best K, Beaulieu-Bonneau S, et al. Physical and cognitive impairments in people suffering from long COVID: protocol for a longitudinal population-based cohort study. BMJ Open 2023;13(3):e064054 [FREE Full text] [doi: 10.1136/bmjopen-2022-064054] [Medline: 36921943]
62. Radin JM, Vogel JM, Delgado F, Coughlin E, Gadaleta M, Pandit JA, et al. Long-term changes in wearable sensor data in people with and without long COVID. NPJ Digit Med 2024;7(1):246 [FREE Full text] [doi: 10.1038/s41746-024-01238-x] [Medline: 39271927]
63. Stewart C, Ranjan Y, Conde P, Sun S, Zhang Y, Rashid Z, et al. Physiological presentation and risk factors of long COVID in the UK using smartphones and wearable devices: a longitudinal, citizen science, case-control study. Lancet Digit Health 2024;6(9):e640-e650 [FREE Full text] [doi: 10.1016/S2589-7500(24)00140-7] [Medline: 39138096]

64. Parker M, Sawant HB, Flannery T, Tarrant R, Shardha J, Bannister R, et al. Effect of using a structured pacing protocol on post-exertional symptom exacerbation and health status in a longitudinal cohort with the post-COVID-19 syndrome. J Med Virol 2023;95(1):e28373 [FREE Full text] [doi: 10.1002/jmv.28373] [Medline: 36461167]
65. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]

Abbreviations

CDC: Centers for Disease Control and Prevention
FHCSD: Family Health Centers of San Diego
FQHC: Federally Qualified Health Center
GAD-7: Generalized Anxiety Disorder-7
GMM: growth mixture modeling
HR: heart rate
HRV: heart rate variability
IACC: infection-associated chronic condition
LC&FIRP: Long COVID and Fatiguing Illness Recovery Program
LME: linear mixed effect
ME/CFS: myalgic encephalomyelitis/chronic fatigue syndrome
MET: metabolic equivalent
MVPA: moderate to vigorous physical activity
PEM: post-exertional malaise
PHQ-2: Patient Health Questionnaire-2
PRO: patient-reported outcome
PROMIS: Patient-Reported Outcomes Measurement Information System
PROMIS-29: Patient-Reported Outcomes Measurement Information System-29
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
teleECHO: Extension for Community Healthcare Outcomes
WHO: World Health Organization

Edited by S Brini; submitted 16.05.25; peer-reviewed by J Ho, H Yu; comments to author 01.09.25; accepted 20.11.25; published 31.12.25.

Please cite as:
Zhang TM, Sharp SP, Scott JD, Taren D, Samaniego JC, Unger ER, Bertolli J, Lin JMS, Ramers CB, Godino JG
Characterization of Post-Viral Infection Behaviors Among Patients With Long COVID: Prospective, Observational, Longitudinal Cohort Analyses of Fitbit Data and Patient-Reported Outcomes
JMIR Form Res 2025;9:e77644
URL: https://formative.jmir.org/2025/1/e77644
doi:10.2196/77644
PMID:

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org