

Artificial Intelligence and Digital Patient Feedback-Supported Cognitive Behavioral Therapy for Chronic Pain: Do Patients with Socioeconomic Vulnerabilities Benefit?

Peter Nilsson^{1*}, Eva Johansson¹, Lars Andersson²

¹ Department of Pain Medicine and Digital Therapeutics, Faculty of Medicine, Karolinska Institute, Stockholm, Sweden.

² Department of Health Equity and Behavioral Health, Faculty of Medicine, Lund University, Lund, Sweden.

*E-mail ✉ peter.nilsson@ki.se

Abstract

A recent comparative effectiveness study found that individuals with chronic pain who received cognitive behavioral therapy augmented by artificial intelligence and digital feedback (AI-CBT-CP) were more likely to achieve clinically meaningful gains in pain-related disability and pain intensity at the six-month mark compared to those assigned to conventional telephone-based CBT-CP. Questions remain regarding how AI and digital approaches affect socially disadvantaged populations. This analysis explored differences in full treatment session completion rates and clinically significant improvements in pain-related disability and intensity across patient subgroups defined by social determinants of health (SDOH). Social determinants of health factors examined included age, race, gender, education level, income, marital status, geographic accessibility, and clinical severity. Multivariate models incorporating interaction terms evaluated whether these SDOH factors moderated treatment engagement and response to AI-CBT-CP relative to standard telephone CBT-CP. Approximately half of the participants (52.9%) were aged 65 years or older, 10.8% were women, and 19.1% identified as Black or as having multiple racial backgrounds. AI-CBT-CP was associated with relatively strong session completion rates across all SDOH subgroups, with no subgroup showing higher completion rates under standard telephone CBT-CP. The superior effects of AI-CBT-CP on pain-related disability and intensity were largely consistent across SDOH categories. AI-CBT-CP produced a stronger relative benefit for pain-related disability among patients under 65 years old ($P = .002$). In no SDOH subgroup did standard telephone CBT-CP outperform AI-CBT-CP on either disability or pain intensity outcomes. The results indicate that patients facing SDOH-related disadvantages do not experience reduced engagement or inferior outcomes with AI- and digitally supported CBT-CP compared to traditional telephone CBT-CP. AI-CBT-CP appears capable of addressing access barriers without increasing disparities, offering particular value to underserved groups living with chronic pain.

Keywords: Artificial intelligence, Cognitive behavioral therapy, Chronic pain, Digital health, Social determinants of health, Health disparities

Introduction

Cognitive behavioral therapy (CBT) represents a safe and well-supported intervention for various mental and physical conditions, including chronic pain [1]. CBT for

chronic pain (CBT-CP) generally involves 6–12 weekly therapist-led sessions that address unhelpful thinking patterns and encourage helpful behavioral changes. Due to limited therapist availability, the demands of repeated in-person visits, and inadequate reimbursement, many patients lack access to CBT-CP or receive fewer sessions than recommended. Telephone-based CBT-CP and digitally delivered therapeutic content can enhance participation while lowering costs [2]. Nevertheless, access to and familiarity with technology remain significant challenges [3, 4]. Among patients at elevated risk for poor engagement and results because of social determinants of health (SDOH), expanded reliance on digitally delivered CBT-CP without careful monitoring

Access this article online

<https://smerpub.com/>

Received: 15 April 2025; Accepted: 29 June 2025

Copyright CC BY-NC-SA 4.0

How to cite this article: Nilsson P, Johansson E, Andersson L. Artificial Intelligence and Digital Patient Feedback-Supported Cognitive Behavioral Therapy for Chronic Pain: Do Patients with Socioeconomic Vulnerabilities Benefit? *J Med Sci Interdiscip Res.* 2025;5(1):226-36. <https://doi.org/10.51847/thMvHLkqoh>

of subgroup differences could worsen existing inequities [5]. In particular, AI-supported digital interventions have raised concerns that algorithmic biases or non-representative training data might disadvantage underrepresented patient groups—who may have distinct needs [6, 7].

In a recently published randomized trial ($n = 278$), we tested reinforcement learning—an AI approach—to improve CBT-CP efficiency. The system automatically tailored session types and lengths according to patients' daily interactive voice response (IVR) feedback on pedometer-recorded step counts and pain interference [8]. A higher percentage of participants in the AI-CBT-CP arm finished all treatment weeks than those in the standard arm (81.6% versus 57.3%, $P < .001$). Relative to conventional telephone CBT-CP, the AI-supported version produced non-inferior gains in pain-related disability at three months (the primary endpoint) while requiring less than half the therapist time. At six months, a significantly larger share of AI-CBT-CP patients showed clinically meaningful improvements in pain-related disability [9] compared with standard CBT-CP (37% versus 19%, $P < .01$). Similarly, more AI-CBT-CP participants achieved clinically meaningful reductions in pain intensity on the Numerical Rating Scale (29% versus 17%, $P = .03$) [10].

Although promising, these outcomes warrant careful consideration before widespread implementation. Average treatment effects reported in trials can obscure smaller benefits—or even potential harms—in specific patient subgroups [11]. Several SDOH-related factors may affect patients' capacity and willingness to use digital self-management tools. For instance, older individuals or those with lower educational attainment might show less enthusiasm for digital CBT or require additional instruction and support [12]. Within the Department of Veterans Affairs (VA) system, where the trial took place [8], women constitute a smaller patient segment, and their VA care experiences could shape their involvement with and gains from digital programs [13]. AI-driven interactions might inadvertently disadvantage women if their patterns differ meaningfully from the predominantly male data used to develop models [14, 15]. Failing to analyze outcomes by gender could overlook such variations.

Engagement and treatment response may also differ by race [4]. Research indicates notable differences across racial groups in pain intensity, functional limitations, and associated socioeconomic difficulties [16, 17].

Telephone or remote CBT-CP could particularly help racial minority patients who encounter obstacles to in-person services. At the same time, specialists emphasize the need to verify benefits within racial subgroups to prevent reinforcing inequities, given possible variations in engagement and the influence of AI systems [6, 7, 18]. Other SDOH elements linked to lower engagement or results—such as social support, depressive symptoms, and baseline pain severity—could likewise affect the comparative value of AI-CBT-CP versus standard telephone psychotherapy [19, 20].

The goal of this secondary analysis was to assess whether the favorable engagement and effectiveness findings for AI-CBT-CP observed in the parent trial [8] held steady across patient subgroups defined by SDOH indicators. We first assessed completion of all treatment weeks among subgroups potentially facing greater access hurdles. We then investigated differences in treatment benefit, concentrating on the two pre-specified secondary outcomes—clinically meaningful improvement in pain-related disability and pain intensity at six months—where AI-CBT-CP showed statistically significant advantages over standard care in the full sample.

Materials and Methods

The parent study's design has been detailed previously [8, 21]. Briefly, it was a randomized, non-inferiority comparative effectiveness trial involving 278 patients with chronic back pain recruited from two VA healthcare facilities. All participants were offered 10 weeks of telephone-based CBT-CP. In the AI-CBT-CP arm, daily IVR patient feedback informed an AI system that recommended weekly treatment formats: a 45-minute therapist telephone session, a 15-minute therapist telephone session, or an IVR-based personalized therapist message. The standard CBT-CP arm received ten 45-minute therapist-delivered telephone sessions. Patient-reported measures were collected at three and six months after randomization. The primary outcome was the Roland Morris Disability Questionnaire (RMDQ) score at three months.

Measurement

The primary focus was on two key metrics: the percentage of patients who finished all 10 scheduled treatment weeks and the percentage in each study arm who achieved a clinically meaningful improvement at the six-month follow-up in pain-related disability and pain

intensity. Treatment engagement duration (number of weeks) was determined from automatically recorded intervention logs and therapist documentation. The selected pain assessment tools followed recommendations from the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) [10, 22]. Pain-related disability was evaluated with the Roland Morris Disability Questionnaire (RMDQ), a validated instrument for individuals experiencing chronic non-cancer pain [9, 22]. Pain intensity was measured using the Numeric Rating Scale (NRS) [10]. For both the RMDQ and NRS, a clinically meaningful change at six months was operationalized as a 30% reduction from each patient's baseline score [10].

SDOH variables were derived from existing study data and aligned with established domains suggested for routine screening and documentation in electronic health records [23, 24]. Age was grouped into < 65 years versus 65 years and older. Gender was based on self-report (male versus female). Race was classified as White versus Black/African-American, multiple racial identities, unknown, or refused. Among the 53 patients categorized as non-White, 67.9% (n=36) identified as Black/African-American and 26.4% (n = 14) reported multiple racial identities, with three patients having unknown race. Education was dichotomized as high school diploma or less versus any education beyond high school. Annual household income was split at <\$30,000 versus ≥\$30,000. Marital status or living in a long-term relationship served as a proxy for social support. Geographic access was assessed by one-way travel time to the nearest VA facility (<1 hour versus ≥1 hour). Baseline pain interference was quantified using the seven-item interference subscale of the Brief Pain Inventory (BPI-I) [10], with a score of 5 or higher indicating at least moderate interference [25, 26]. Depressive symptom severity was measured with the Patient Health Questionnaire-9 (PHQ-9), where a score of 10 or higher indicated at least moderate depression [25].

Statistical analysis

We began by summarizing the overall sample characteristics and then by randomization arm across all SDOH indicators. Next, we calculated the proportion of patients in each arm who completed the full 10 treatment weeks. These comparisons involved unadjusted chi-square tests of completion rates within each SDOH-

defined stratum, followed by adjusted odds ratios for full session completion by randomization group within each stratum. Adjustments mirrored the primary trial analyses [8] and included age and recruitment site. To test for differential engagement by SDOH subgroup, we used logistic regression models containing main effects for treatment arm and SDOH indicator plus their interaction term.

We performed parallel analyses for the two binary six-month outcomes (clinically meaningful improvement on the RMDQ and NRS). We fitted generalized linear mixed models (GLMMs) with a random intercept to model each outcome at the 3- and 6-month time points. Predictors included randomization group, time, the SDOH indicator, and interaction terms (group × SDOH, group × time, SDOH × time, and the three-way group × SDOH × time interaction). All models were adjusted for the continuous baseline value of the respective outcome (RMDQ or NRS), age, and recruitment site. When age served as the SDOH variable, it was not included as a continuous covariate to prevent collinearity. Because small cell sizes caused convergence problems in models examining RMDQ response by age and gender subgroups, Firth logistic regression was used instead of GLMM [27]. In the model for NRS improvement that included BPI-I as the SDOH indicator, baseline NRS was omitted as a covariate because both measures assess overlapping aspects of pain.

Role of the funding source

The funding organizations had no involvement in the study design or conduct; data collection, management, analysis, or interpretation; manuscript preparation, review, or approval; or the decision to submit the manuscript for publication.

Results and Discussion

Patient characteristics at baseline

The total sample (n = 278) consisted of roughly balanced proportions of patients younger than 65 years (47.1%) and those aged 65 or older (52.9%); (**Table 1**). The large majority were men (89.2%), with women comprising 10.8% of participants. Most identified as White (80.9%), while 19.1% (53 patients) reported Black race, multiple racial identities, unknown race, or refused (three patients). Over one-third (35.5%) were not married or in a long-term relationship, 29.4% reported household incomes below \$30,000 per year, and 27.7% needed one

hour or more to travel to VA care. At baseline, 63.8% of patients experienced at least moderate pain-related interference on the BPI-I, and 24.8% had PHQ-9 scores indicating at least moderate depressive symptoms. While

minor differences existed in the distribution of these characteristics between randomization arms, the groups were largely comparable on the SDOH-related factors.

Table 1. Baseline characteristics.

Empty cell	Total N (%)	AI-CBT-CP N (%)	Phone CBT-CP N (%)
All patients	278 (100.0)	168 (60.4)	110 (39.6)
Age < 65 years	131 (47.1)	83 (49.4)	48 (43.6)
Age 65+ years	147 (52.9)	85 (50.6)	62 (56.4)
Male	248 (89.2)	147 (87.5)	101 (91.8)
Female	30 (10.8)	21 (12.5)	9 (8.2)
White race	225 (80.9)	133 (79.2)	92 (83.6)
Other race ^a	53 (19.1)	35 (20.8)	18 (16.4)
High school or less ^b	56 (20.3)	29 (17.4)	27 (24.8)
Some college+	220 (79.1)	138 (82.6)	82 (75.2)
Not married ^c	98 (35.5)	59 (35.3)	39 (35.8)
Married	178 (64.5)	108 (64.7)	70 (64.2)
Income < \$30,000 ^d	78 (29.4)	40 (25.0)	38 (36.2)
Income ≥ \$30,000	187 (70.6)	120 (75.0)	67 (63.8)
Travel < 1 hour	201 (72.3)	114 (67.9)	87 (79.1)
Travel 1+ hours	77 (27.7)	54 (32.1)	23 (20.9)
BPI-I ^e < 5	101 (36.3)	64 (38.1)	37 (33.6)
BPI-I 5+	177 (63.8)	104 (61.9)	73 (66.4)
PHQ ^f < 10	209 (75.2)	123 (73.2)	86 (78.2)
PHQ 10+	69 (24.8)	45 (26.8)	24 (21.8)

^a Of the non-White participants, 72% described their race as Black or African-American, with the other 28% indicating multiple racial backgrounds or declining to specify/unknown.

^b N = 276.

^c N = 276.

^d N = 265.

^e BPI-I refers to the average across the seven interference items from the Brief Pain Inventory. A score of 5 or above on this scale denotes at least moderate levels of pain-related interference.

^f PHQ stands for the Patient Health Questionnaire. A score of 10 or higher corresponds to at least moderate depressive symptoms.

Session engagement

Consistent with earlier findings [8], a substantially higher percentage of patients assigned to AI-CBT-CP completed all 10 treatment weeks than those assigned to standard CBT-CP (81.6% versus 57.3%, $P < .001$). This pattern of increased treatment participation in the AI-CBT-CP arm relative to the CBT-CP arm held across every SDOH

subgroup we had identified as potentially vulnerable to difficulties with digital intervention use (**Table 2**).

In unadjusted comparisons, women in the AI-CBT-CP group showed markedly higher completion rates (81.0%) than women in the CBT-CP group (44.4%, $P = .046$). Similarly, individuals who identified as Black or as having multiple racial identities were significantly more

likely to complete the full course of sessions when receiving AI-CBT-CP than when receiving standard CBT-CP ($P = .002$). The narrowest gaps in engagement between the two arms appeared among patients with annual household incomes below \$30,000 ($P = .373$) and those who lived 1+ hours away from VA facilities ($P = .56$). Nevertheless, even within these subgroups, the adjusted odds ratios pointed toward greater engagement under the AI-supported approach.

Formal tests of moderation using multivariate models that included interaction terms between each SDOH indicator and randomization group revealed no statistically significant interactions (all $P \geq .05$). Therefore, the overall advantage in treatment completion observed with AI-CBT-CP did not differ meaningfully across SDOH subgroups.

Table 2. Proportion of patients completing all 10 weekly sessions.

Empty cell	AI-CBT-CP		Phone CBT-CP		P-value	AOR ^a (95 % CI)	α P-value ^b
	Total N	%, N	Total N	%, N			
All patients	168	81.6%, 137	110	57.3%, 63	< .001	3.5 (2.0, 6.0)	
Age < 65 years	83	80.7%, 67	48	50.0%, 24	< .001	4.3 (1.9, 9.5)	.60
Age 65+ years	85	82.4%, 70	62	62.9%, 39	.008	2.9 (1.3, 6.4)	
Male	147	81.6%, 120	101	58.4%, 59	< .001	3.4 (1.9, 6.1)	.53
Female	21	81.0%, 17	9	44.4%, 4	.046	4.9 (.9, 28.2)	
White race	133	80.5%, 107	92	59.8%, 55	.001	2.9 (1.6, 5.4)	.22
Other race ^c	35	85.7%, 30	18	44.4%, 8	.002	7.4 (1.9, 28.5)	
High school or less	29	75.9%, 22	27	59.3%, 16	.184	2.6 (.7, 9.2)	.24
Some college+	138	83.3%, 115	82	47.6%, 39	< .001	4.1 (2.1, 7.7)	
Not married	59	77.9%, 46	39	61.5%, 24	.078	2.1 (.8, 5.2)	.34
Married	108	84.3%, 91	70	55.7%, 39	< .001	4.6 (2.2, 9.4)	
Income < \$30,000	40	75.0%, 30	38	65.8%, 25	.373	1.8 (.6, 5.0)	.21
Income $\geq$ \$30,000	120	84.2%, 101	67	52.2%, 35	< .001	5.0 (2.5, 10.1)	
Travel < 1 hour	114	84.2%, 96	87	54.0%, 47	< .001	5.5 (2.7, 11.1)	.31
Travel 1+ hours	54	75.9%, 41	23	69.6%, 16	.560	1.6 (.5, 4.9)	
BPI-I ^d < 5	64	85.9%, 55	37	62.2%, 23	.006	4.1 (1.5, 11.5)	.30
BPI-I 5+	104	78.9%, 82	73	54.8%, 40	< .001	3.6 (1.8, 7.2)	
PHQ ^e < 10	123	82.9%, 102	86	57.0%, 49	< .001	4.1 (2.1, 7.8)	.57
PHQ 10+	45	77.8%, 35	24	58.3%, 14	.090	2.7 (.9, 8.23)	

^a Adjusted odds ratios (AORs) calculated separately within each subgroup (refer to Methods section for further information).

^b P-value associated with the interaction term between the SDOH moderator and treatment assignment (refer to Methods section for further information).

^c Combines participants identifying as Black/African-American, those with multiple racial identities, and those with unknown or refused responses.

^d BPI-I represents the mean score derived from the seven interference items of the Brief Pain Inventory. Scores reaching 5 or higher signify at least moderate pain-related interference.

^e PHQ stands for Patient Health Questionnaire. Scores of 10 or above correspond to at least moderate depressive symptoms.

Although limited by sample size, exploratory within-arm analyses offer additional perspective on potential engagement differences by SDOH factors by examining

completion rates across categories of each indicator separately for each treatment condition.

For example, women assigned to conventional telephone CBT-CP showed modestly lower full-session completion rates than men in the same condition (44.4% vs. 58.4%). In sharp contrast, completion rates were comparably high for both women and men who received AI-CBT-CP (81.0% and 81.6%, respectively). Similarly, participants identifying as Black or reporting multiple racial identities had lower completion rates than White participants when assigned to standard CBT-CP (44.4% vs. 59.8%). Still, this gap disappeared in the AI-supported arm (85.7% vs. 80.5%). Within the standard CBT-CP group, those facing longer travel times (1+ hours) to VA facilities actually completed all sessions at higher rates than those with shorter commutes (69.6% vs. 54.0%). However, the opposite pattern appeared in the AI-CBT-CP group, where patients with longer travel times had slightly lower completion rates (75.9%) than those living closer (84.2%).

Clinically meaningful improvements in pain-related outcomes

Consistent with the main trial report [8], participants randomized to AI-CBT-CP were more likely to demonstrate a clinically meaningful gain on the RMDQ at the six-month assessment than those in the standard

CBT-CP arm (36.7% vs. 19.0%, $P < .001$). This comparative advantage appeared in several—but not all—SDOH-defined subgroups (**Table 3**). The benefit was especially pronounced among patients younger than 65: 46.4% in the AI-CBT-CP condition achieved meaningful RMDQ improvement, compared with only 7.3% in the standard arm ($P < .001$). No difference between arms was seen among patients aged 65 and older (27.1% vs. 27.8%, $P = .95$).

For men, the AI-CBT-CP arm produced significantly higher rates of meaningful RMDQ improvement than the CBT-CP arm (36.7% vs. 20.7%, $P = .02$). Among women, 36.8% in the AI-CBT-CP group reached this threshold, while zero women in the standard CBT-CP group did so ($P = .16$). Subgroup bivariate tests indicated statistically significant advantages for AI-CBT-CP among White patients; differences did not attain significance among Black/African-American patients or those with multiple racial identities. Clear benefits favoring AI-CBT-CP emerged across categories of marital status, household income, and baseline pain interference on the BPI-I (all $P \leq .05$). Comparable advantages were also noted for patients with shorter travel distances to VA care and those with below-moderate depressive symptoms.

Table 3. Clinically significant improvement^a in roland morris disability questionnaire (RMDQ) scores between baseline and six months.

Empty cell	AI-CBT-CP		Phone CBT-CP		P-value	AOR ^b (95 % CI)	α P-value ^c
	Total N	%, N	Total N	%, N			
All patients	139	36.7%, 51	95	19.0%, 18	< .01	4.7 (1.4, 15.9)	
Age <65	69	46.4%, 32	41	7.3%, 3	< .001	9.2 (2.8, 30.7)	.002
Age 65+	70	27.1%, 19	54	27.8%, 15	.95	1.0 (.4, 2.2)	
Male	120	36.7%, 44	87	20.7%, 18	.02	2.1 (1.1, 4.1)	.37
Female	19	36.8%, 7	8	0%, 0	.16	9.5 (.4, 222.7)	
White race	111	36.9%, 41	81	18.5%, 15	.02	4.6 (1.3, 17.3)	.84
Other race ^d	27	33.3%, 9	13	23.1%, 3	.43	3.3 (.2, 66.0)	
High school or less	23	34.8%, 8	21	9.5%, 2	.09	16.2 (.7, 389.8)	.39
Some college+	116	37.1%, 43	73	21.9%, 16	.06	3.7 (1.0, 13.8)	
Not married	44	45.5%, 20	34	23.5%, 8	.05	5.2 (.7, 37.0)	.91
Married	95	32.6%, 31	60	16.7%, 10	.03	4.5 (1.0, 20.4)	
Income < \$30,000	27	33.3%, 9	32	9.4%, 3	.02	15.5, (1.0, 20.4)	.41
Income ≥ \$30,000	106	38.7%, 41	59	22.0%, 13	.03	4.3 (.9, 20.1)	

Travel < 1 hour	93	38.7%, 36	76	17.1%, 13	.002	6.5 (1.6, 27.3)	.31
Travel 1+ hours	46	32.6%, 15	19	26.3 %, 5	.62	1.6 (.2, 16.5)	
BPI-I ^e < 5	58	48.3%, 28	33	27.3%, 9	.050	4.8 (1.0, 24.0)	.78
BPI-I 5+	81	28.4%, 23	62	14.5%, 9	.048	3.5 (.6, 20.6)	
PHQ ^f < 10	104	35.6%, 37	74	18.9%, 14	.04	4.2 (1.1, 16.3)	.73
PHQ 10+	35	40.0%, 14	21	19.1%, 4	.14	6.8 (.5, 86.8)	

^a A clinically meaningful improvement was operationalized following established pain research recommendations—specifically, a 30% reduction in scores relative to each participant's baseline value.

^b Adjusted odds ratios (AORs) were computed separately within each subgroup stratum (details provided in the Methods section).

^c P-value corresponding to the interaction between the SDOH moderator variable and randomization arm (details provided in the Methods section).

^d Encompasses individuals identifying as Black/African-American, those reporting multiple racial identities, and those with unknown or refused race information.

^e BPI-I denotes the average across the seven interference items from the Brief Pain Inventory. A score of 5 or greater signals at least moderate pain-related interference.

^f PHQ refers to the Patient Health Questionnaire. Scores of 10 or higher reflect at least moderate depressive symptoms.

In no SDOH subgroup did either the unadjusted comparisons or the adjusted odds ratios (AORs) from the multivariate models indicate that clinically meaningful gains in RMDQ scores occurred more frequently with standard CBT-CP than with AI-CBT-CP. When interaction terms between SDOH indicators and treatment assignment were formally tested, only the term involving age was statistically significant.

For the outcome of clinically meaningful improvement in NRS pain intensity scores, no SDOH subgroup showed a reversal in which standard CBT-CP outperformed AI-CBT-CP (**Table 4**). Within-subgroup bivariate tests revealed statistically significant advantages for AI-CBT-

CP among younger patients, men, those who were married or in a long-term partnership, and individuals with higher baseline BPI-I interference levels (all $P \leq .05$). Multivariate models examining interaction terms found no statistically significant differences in the magnitude of NRS benefit across any SDOH subgroups. That said, the adjusted odds ratio point estimates hinted that AI-CBT-CP may have produced stronger effects among patients who were under 65 years old, possessed a high school education or less, reported household incomes below \$30,000, required 1+ hours of travel time to VA facilities, and exhibited at least moderate depressive symptoms.

Table 4. Clinically meaningful improvement^a on the Numeric Rating Scale (NRS) for pain intensity from baseline to six-month follow-up.

Empty cell	AI-CBT-CP		Phone CBT-CP		P-value	AOR ^b (95 % CI)	α P-value ^c
	Total N	%, N	Total N	%, N			
All patients	138	29.0%, 40	94	17.0%, 16	.03	4.5 (1.1, 18.1)	
Age < 65	68	30.9%, 21	40	10.0%, 4	.04	11.3 (1.2, 108.8)	.20
Age 65+	70	27.1%, 19	54	22.2%, 12	.54	1.7 (.3, 10.1)	
Male	119	30.3%, 36	86	17.4%, 15	.04	4.7 (1.1, 20.2)	.88
Female	19	21.1%, 4	8	12.5%, 1	.59	3.3 (.04, 274.9)	
White race	111	26.1%, 29	80	16.3%, 13	.07	4.3 (.9, 21.0)	.96
Other race ^d	26	38.5%, 10	13	23.1%, 3	.37	4.8 (.2, 147.5)	
High school or less	22	31.8%, 7	20	20.0%, 4	.16	9.6 (.4, 230.4)	.62
Some college+	116	28.5%, 33	73	16.4%, 12	.08	4.0 (.8, 18.9)	
Not married	45	33.3%, 15	34	23.5%, 8	.34	2.8 (.3, 26.5)	.50

Married	93	26.9%, 25	59	13.6%, 8	.05	7.5 (1.2, 48.2)	
Income < \$30,000	28	28.6%, 8	32	12.5%, 4	.12	12.7 (.8, 213.5)	.36
Income ≥ \$30,000	105	29.5%, 31	58	20.7%, 12	.22	2.9 (.6, 14.4)	
Travel < 1 hour	92	26.1%, 24	76	18.4%, 14	.24	2.5 (.5, 12.4)	.20
Travel 1+ hours	46	34.8%, 16	18	11.1%, 2	.06	24.9 (1.0, 641.5)	
BPI-I ^e < 5	56	32.1%, 18	33	24.2%, 8	.43	3.6 (.6, 21.4)	.89
BPI-I 5+	82	26.8%, 22	61	13.1%, 8	.046	4.4 (.5, 36.6)	
PHQ ^f < 10	104	28.9%, 30	73	17.8%, 13	.09	3.7 (.8, 17.0)	.54
PHQ 10+	34	29.4%, 10	21	14.3%, 3	.14	10.6 (.5, 240.0)	

Note: The table columns have been formatted with sufficient horizontal spacing to prevent wrapping. Each row's content remains on a single line for clarity in preview and display.

^a Clinically meaningful improvement was determined according to standards endorsed in pain research recommendations, defined as a 30% reduction from baseline scores.

^b Adjusted odds ratios (AORs) were derived separately for each subgroup (additional information is available in the Methods).

^c P-value reflecting the interaction effect between the SDOH moderator and treatment randomization arm (additional information is available in the Methods).

^d Groups together participants who identified as Black/African-American, those selecting multiple racial identities, and individuals with unknown or refused racial data.

^e BPI-I is the mean value calculated from the seven interference items of the Brief Pain Inventory. Scores at or above 5 reflect at least moderate pain-related interference.

^f PHQ refers to the Patient Health Questionnaire. A score of 10 or higher signals at least moderate depressive symptoms.

Scholars, health system leaders, and patient representatives frequently stress the risks of generalizing overall trial findings to diverse patient populations without carefully assessing how effects might differ by social determinants of health (SDOH). Experts in care delivery and policy have specifically raised alarms that digital tools, especially those incorporating artificial intelligence, could unintentionally create or widen gaps in outcomes across sociodemographic and clinical categories. In response to these valid issues, trial data should be scrutinized to detect any meaningful differences in patient participation and therapeutic gains across SDOH-defined groups.

The current subgroup analyses indicate that the previously documented overall superiority of AI-CBT-CP extended to most SDOH categories. Although the strength of this advantage varied somewhat, patients assigned to AI-CBT-CP generally showed higher rates of completing the full set of treatment weeks compared with those receiving standard telephone CBT-CP across nearly all subgroups. This pattern is particularly promising given existing evidence on the digital divide and concerns about lower participation among traditionally underserved individuals.

Because having telephone access was required for trial enrollment, the findings do not address obstacles faced by people lacking reliable phone service or internet

connectivity. Nevertheless, recent national statistics indicate that such barriers are diminishing. By 2021, cell phone ownership reached 98% among adult women, 99% among Black Americans, 96% among those with no more than a high school education, and 97% among adults earning less than \$30,000 annually [28]. Even for interventions requiring internet access, most disadvantaged groups could participate (for example, 85% of Black Americans and 76% of those with incomes below \$30,000 owned smartphones that year). While individuals without any telephone access remain an important group to consider, telehealth and digitally assisted disease management strategies represent a meaningful step toward broadening access to services like CBT for chronic pain among disadvantaged populations.

Regarding clinically important treatment effects, the benefits of AI-CBT-CP seen in the full sample largely appeared across most (though not every) SDOH subgroup. Notably, patients younger than 65 were far more likely to achieve a meaningful RMDQ improvement with AI-CBT-CP than with standard care (46% versus 7%). Among those aged 65 and older, the likelihood of benefit was essentially identical regardless of which intervention was received. This observation is noteworthy because older adults in the AI-CBT-CP arm still completed all sessions at higher rates than their

counterparts in standard CBT-CP and matched the completion rates of younger patients in the AI-supported condition.

Given ongoing concerns that AI technologies might maintain or intensify healthcare inequities [7, 18], it is encouraging that this program delivered comparable advantages across groups defined by key SDOH characteristics. Progress is being made in developing fairness-aware algorithms designed to advance equity objectives and deliver appropriate care to diverse patients. AI systems can be engineered with multiple priorities in mind, such as directing resources toward those most likely to gain from treatment or minimizing racial and gender biases [29]. As adoption of AI-supported healthcare tools grows, continued advances in equitable algorithm design will be essential to maximize benefits while minimizing risks.

Women and men achieved nearly identical rates of clinically meaningful RMDQ improvement at six months within the AI-CBT-CP group (37% for both). In contrast, no women assigned to standard CBT-CP reached this level of improvement. Additionally, patients with a high school education or less had substantially greater chances of meaningful RMDQ gains with AI-CBT-CP than with standard treatment. Stratified results also pointed to strong relative benefits for AI-CBT-CP among those starting with at least moderate pain interference. The advantage over standard care appeared at least as large, if not larger, for individuals experiencing significant depressive symptoms.

Participants in this research study faced no out-of-pocket expenses for either AI-CBT-CP or the comparison condition. Moreover, because the trial occurred within the VA system, most (though not all) patients would have faced minimal or no copayments if the program were rolled out clinically. Copayment structures differ considerably outside the VA in public and private settings, potentially affecting willingness to participate for some individuals. Importantly, however, AI-CBT-CP produced outcomes equivalent to a full 10-session therapist-delivered course while requiring only 12.3% of the standard 45-minute therapist sessions—roughly one such session per patient on average [21]. Copayments for the shorter AI-CBT-CP session options (15-minute therapist calls or IVR messages) are unknown but would presumably be lower than those for full sessions. Consequently, although non-VA patients might encounter some financial hurdles, the overall costs associated with AI-CBT-CP would likely remain far

lower than those of conventional programs, assuming such services are even available.

This investigation had limited statistical power to detect interactions between SDOH factors and treatment assignment for either engagement or clinical outcomes. The modest number of female participants calls for larger future studies focused on gender differences, especially considering the patterns observed here. Due to small numbers within specific racial categories, non-White patients were grouped. While having a partner can confer benefits, this dataset did not measure other aspects of social support [30]. Finally, several critical SDOH elements—such as housing instability, food insecurity, health insurance status, or experiences of discrimination—could not be examined.

Notwithstanding these constraints, the results provide no evidence that priority populations, including women, patients identifying as Black or other non-White racial groups, and those with lower education or income, experience reduced engagement with AI-CBT-CP relative to standard telephone CBT-CP. Similarly, the data do not support the idea that these groups derive less clinical benefit in terms of meaningful pain improvement when offered the AI-supported intervention. Overall, these findings are reassuring and suggest that digitally assisted approaches can help alleviate, rather than worsen, disparities in access to effective pain CBT.

Conclusion

In conclusion, the subgroup analyses demonstrate that the primary benefits of AI-CBT-CP documented in the main outcomes paper [8] extended broadly to underrepresented VA patients with chronic pain, such as women, individuals from racial groups other than White, and those with lower education or income. The elevated rate of clinically meaningful RMDQ improvements at six months under AI-CBT-CP was driven almost entirely by patients younger than 65; older adults showed roughly comparable success rates in both arms. Innovative digital strategies for delivering sophisticated behavioral treatments like CBT-CP can therefore meaningfully reduce inequities in both access to care and treatment effectiveness.

Acknowledgments: None

Conflict of Interest: None

Financial Support: None

Ethics Statement: None

References

1. Chou R, Huffman LH. Nonpharmacologic therapies for acute and chronic low back pain: a review of the evidence for an American Pain Society/American College of Physicians clinical practice guideline. *Ann Intern Med.* 2007;147(7):492-504.
2. Pimm TJ, Williams LJ, Reay M, Pickering S, Lota R, Coote L. An evaluation of a digital pain management programme: clinical effectiveness and cost savings. *Br J Pain.* 2020;14(4):238-49.
3. Lee M, Kang D, Yoon J, Kim S, Lee J, Park H, et al. The difference in knowledge and attitudes of using mobile health applications between actual users and non-users among adults aged 50 and older. *PLoS One.* 2020;15(10):e0240581.
4. Walker DM, Hefner JL, Fareed N, Huerta TR, McAlearney AS. Exploring the digital divide: age and race disparities in use of an inpatient portal. *Telemed J E Health.* 2020;26(5):603-13.
5. Chang JE, Lai AY, Gupta A, Nguyen AM, Berry CA, Shelley DR. Rapid transition to telehealth and the digital divide: implications for primary care access and equity in a post-COVID era. *Milbank Q.* 2021;99(2):340-68.
6. McCradden MD, Joshi S, Mazwi M, Anderson JA. Ethical limitations of algorithmic fairness solutions in health care machine learning. *Lancet Digit Health.* 2020;2(5):e221-3.
7. Parikh RB, Teeple S, Navathe AS. Addressing bias in artificial intelligence in health care. *JAMA.* 2019;322(24):2377-8.
8. Piette JD, Newman S, Krein SL, Marinec N, Chen J, Williams DA, et al. Patient-centered pain care using artificial intelligence and mobile health tools: a randomized, comparative effectiveness trial. *JAMA Intern Med.* 2022;182(9):975-83.
9. Roland M, Fairbank J. The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. *Spine.* 2000;25(24):3115-24.
10. Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, Farrar JT, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. *J Pain.* 2008;9(2):105-21.
11. Kent DM, Hayward RA. Limitations of applying summary results of clinical trials to individual patients: the need for risk stratification. *JAMA.* 2007;298(10):1209-12.
12. Bhattarai P, Phillips JL. The role of digital health technologies in management of pain in older people: an integrative review. *Arch Gerontol Geriatr.* 2017;68:14-24.
13. Kehle-Forbes SM, Harwood EM, Spont MR, Sayer NA, Gerould H, Murdoch M. Experiences with VHA care: a qualitative study of US women veterans with self-reported trauma histories. *BMC Womens Health.* 2017;17(1):38.
14. Cirillo D, Catuara-Solarz S, Morey C, Guney E, Subirats L, Mellino S, et al. Sex and gender differences and biases in artificial intelligence for biomedicine and healthcare. *NPJ Digit Med.* 2020;3:81.
15. Samulowitz A, Gremyr I, Eriksson E, Hensing G. "Brave men" and "emotional women": a theory-guided literature review on gender bias in health care and gendered norms toward patients with chronic pain. *Pain Res Manag.* 2018;2018:6358624.
16. Anderson KO, Green CR, Payne R. Racial and ethnic disparities in pain: causes and consequences of unequal care. *J Pain.* 2009;10(12):1187-204.
17. Poleshuck EL, Green CR. Socioeconomic disadvantage and pain. *Pain.* 2008;136(3):235-8.
18. Daneshjou R, Smith MP, Sun MD, Rotemberg V, Zou J. Lack of transparency and potential bias in artificial intelligence data sets and algorithms: a scoping review. *JAMA Dermatol.* 2021;157(11):1362-9.
19. Che X, Cash R, Ng SK, Fitzgerald P, Fitzgibbon BM. A systematic review of the processes underlying the main and buffering effect of social support on the experience of pain. *Clin J Pain.* 2018;34(11):1061-76.
20. Broderick JE, Keefe FJ, Schneider S, Junghaenel DU, Schwartz JE, Stone AA. Cognitive behavioral therapy for chronic pain is effective, but for whom? *Pain.* 2016;157(9):2115-23.
21. Piette JD, Newman S, Krein SL, Marinec N, Chen J, Williams DA, et al. Artificial intelligence (AI) to improve chronic pain care: evidence of AI Learning. *Intell Based Med.* 2022;6:100064.
22. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, et al. Core outcome

- measures for chronic pain clinical trials: IMMPACT recommendations. *Pain*. 2005;113(1-2):9-19.
23. Hill-Briggs F, Adler NE, Berkowitz SA, Chin MH, Gary-Webb TL, Navas-Acien A, et al. Social determinants of health and diabetes: a scientific review. *Diabetes Care*. 2021;44(1):258-79.
 24. Braveman P, Egerter S, Williams DR. The social determinants of health: coming of age. *Annu Rev Public Health*. 2011;32:381-98.
 25. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16(9):606-13.
 26. Cleeland CS. The Brief Pain Inventory: User Guide. Houston: MD Anderson Cancer Center; 2009.
 27. Wang X. Firth logistic regression for rare variant association tests. *Front Genet*. 2014;5:187.
 28. Pew Research Center. Mobile technology and home broadband 2021. Washington, DC: Pew Research Center; 2021.
 29. Fang EX, Wang Z, Wang L. Fairness-oriented learning for optimal individualized treatment rules. *J Am Stat Assoc*. 2022;117(540):1-15.
 30. Reese JB, Somers TJ, Keefe FJ, Mosley-Williams A, Lumley MA. Pain and functioning of rheumatoid arthritis patients based on marital status: is a distressed marriage preferable to no marriage? *J Pain*. 2010;11(10):958-64.