Emotional Trajectories in Chronic Illness: A Longitudinal Analysis of 1.28 Million Conversations with an AI Health Companion

Marshall Gould[1] Isaac Tolley[1]

[1]SharedGenes Inc, London, United Kingdom

Abstract

Background: Chronic illness affects over 130 million adults in the United States alone, with depression and anxiety comorbidity rates reaching 30–50%. Despite this burden, access to mental health support remains constrained by clinician shortages, geographic barriers, and prohibitive costs. AI-driven conversational agents offer a scalable pathway to bridge this gap, yet evidence for their effectiveness in chronic illness populations—where emotional needs differ markedly from acute mental health conditions—remains scarce.

Objective: To characterise within-session emotional trajectories and longitudinal sentiment change among users of Juno, a large-language-model–powered health companion designed for people living with chronic illness, and to evaluate whether sustained engagement predicts emotional improvement.

Methods: We conducted a retrospective observational study of 1,281,637 messages (1,036,834 text; 244,878 voice) exchanged by 12,294 active users between August 2025 and March 2026. Sentiment was assessed via a validated five-category classifier applied to 91,895 sentiment-tagged events. Within-session emotional transitions were analysed using Markov transition matrices and χ^2 tests. Longitudinal trajectories were computed for 1,555 users with $\geq$10 sentiment events across $\geq$2 weeks, comparing first-week and last-week mean sentiment. Thematic coding was performed on a stratified subsample of 4,800 messages.

Results: Within sessions, 65.6% of transitions were toward more positive sentiment, yielding a positive-to-negative shift ratio of 2.5:1 ($\chi^2 = 6{,}377$, $p \approx 0$; $N = 10{,}648$ transitions). The most common recovery pathway was frustration→gratitude ($n = 470$), followed by frustration→connection ($n = 342$) and frustration→relief ($n = 271$). Session length predicted emotional recovery ($r = 0.104$, $p = 5.53 \times 10^{-13}$). Longitudinally, 45.2% of users improved versus 21.2% who declined (ratio 2.13:1; mean improvement +9.0 percentage points). A dose–response relationship was observed: 58.6% of heavy users ($\geq$50 events; $n = 307$) improved, with an inverted-U peak at 31–50 events. Thematic analysis revealed a 15% increase in self-advocacy language and an 11% decrease in pain/distress language from Week 1 to Week 4.

Conclusions: To our knowledge, this is the first large-scale evidence that an AI health companion can facilitate meaningful emotional improvement in a chronic illness population. The consistent recovery patterns, dose–response relationship, and cross-condition generalisability suggest that AI conversational agents are a viable complement to traditional mental health services, particularly where access is limited by waitlists, cost, or geography. These findings are urgently needed given the scale of unmet emotional need in chronic illness.

Keywords: chronic illness; artificial intelligence; conversational agent; sentiment analysis; emotional wellbeing; digital health; longitudinal study; large language model

1 Introduction

1.1 The Emotional Burden of Chronic Illness

Chronic illness represents one of the most significant public health challenges of the 21st century. In the United States, approximately 133 million adults—nearly half the population—live with at least one chronic condition, a figure projected to reach 170 million by 2030 [1, 2]. Globally, chronic non-communicable diseases account for 74% of all deaths [3]. Beyond mortality and physical disability, chronic illness imposes a pervasive emotional burden that is frequently underrecognized in clinical practice.

Depression affects 20–30% of individuals with chronic illness, roughly two to three times the prevalence in the general population [4, 5]. Among specific conditions, the figures are stark: up to 40% of fibromyalgia patients meet criteria for major depressive disorder [6], while 50% of people with multiple sclerosis experience clinically significant depression during their lifetime [7]. Anxiety disorders are similarly elevated, particularly in conditions characterised by unpredictable symptom flares such as postural orthostatic tachycardia syndrome (POTS), where prevalence estimates reach 38% [8].

Critically, the emotional distress associated with chronic illness differs qualitatively from primary psychiatric conditions. It is often rooted in grief over lost function, frustration with medical invalidation, the relentless cognitive load of self-management, and the social isolation that accompanies reduced activity levels [9, 10]. These experiences demand support modalities that acknowledge the legitimacy of physical suffering while fostering adaptive coping—a nuance that generic mental health interventions frequently miss [11].

1.2 The Access Gap

Despite the scale of need, access to appropriate mental health support for people with chronic illness remains severely constrained. In the United Kingdom, National Health Service (NHS) waiting times for psychological therapy average 18 weeks, with some regions exceeding 12 months [12]. In the United States, over 150 million people live in federally designated Mental Health Professional Shortage Areas [13]. Even when services are accessible, they may not be adapted for chronic illness: standard cognitive-behavioural therapy (CBT) protocols frequently require modification to address the reality of permanent or progressive illness rather than assumed recovery [14].

The mismatch between need and provision is compounded by the temporal pattern of emotional distress in chronic illness. Symptom flares, medication side effects, and health-related crises do not adhere to business hours. A patient experiencing a 2 AM pain crisis or post-appointment distress after an invalidating clinical encounter has, in most healthcare systems, no accessible professional support [15].

1.3 AI Conversational Agents as a Bridge

The emergence of large language model (LLM)-based conversational agents offers a potential pathway to address this access gap. Unlike earlier rule-based chatbots, modern LLMs demonstrate nuanced language understanding, contextual reasoning, and the capacity for extended, empathetic dialogue [16, 17]. Several studies have demonstrated that LLM-generated responses can match or exceed human clinician responses in perceived empathy and quality [18, 19].

A growing body of evidence supports the efficacy of AI-driven mental health interventions. Woebot, a CBT-based chatbot, demonstrated significant reductions in depression symptoms over two weeks in a randomised controlled trial [20]. Wysa showed comparable efficacy to human coaching for depression and anxiety in a workplace setting [21]. More recently, Replika users reported decreased loneliness and increased emotional awareness through sustained conversational engagement [22].

However, the existing evidence base has three critical limitations. First, most studies focus on primary mental health conditions (depression, anxiety) rather than the distinct emotional needs of chronic physical illness. Second, sample sizes rarely exceed a few hundred participants, limiting generalisability. Third, longitudinal data beyond 4–8 weeks is scarce, precluding analysis of sustained emotional trajectories [23, 24].

1.4 Juno: An AI Health Companion for Chronic Illness

Juno is a mobile-first AI health companion specifically designed for people living with chronic illness. Built on Anthropic's Claude large language model, Juno integrates 15 therapeutic frameworks—including dialectical behaviour therapy (DBT), acceptance and commitment therapy (ACT), motivational interviewing (MI), and spoon theory—into a conversational interface available 24/7 via text and voice modalities. The system incorporates persistent memory across sessions, health metric tracking, activity pacing tools, and condition-specific knowledge spanning over 200 chronic conditions.

Juno's design reflects a fundamental insight from the chronic illness literature: that emotional support for this population must validate physical reality while fostering agency, rather than pursuing symptom elimination [25]. This means that for a person with a degenerative condition, *stability* in emotional wellbeing is itself a positive outcome, and that the pathway from frustration to relief (rather than forced positivity) represents an authentic therapeutic trajectory.

1.5 Study Objectives

This study presents the first large-scale analysis of emotional trajectories in conversations between people with chronic illness and an AI health companion. Our objectives are:

1. To characterise within-session emotional transitions and identify predominant recovery pathways.
2. To assess longitudinal changes in sentiment over weeks to months of engagement.
3. To evaluate dose–response relationships between engagement intensity and emotional improvement.
4. To examine whether emotional trajectories differ by chronic condition.
5. To identify thematic shifts in user language that may reflect psychological growth beyond sentiment alone.

2 Theoretical Framework

Juno's conversational design draws on 15 established therapeutic and theoretical frameworks, integrated dynamically based on user context, emotional state, and condition. This section describes each framework and its application within the system.

2.1 Dialectical Behaviour Therapy (DBT)

DBT, developed by Linehan [26], provides the primary scaffolding for emotional regulation within Juno. The system implements six levels of validation, ranging from active listening (Level 1) through radical genuineness (Level 6). In chronic illness contexts, DBT's dialectical stance—holding that a user's suffering is real *and* that adaptive change is possible—is particularly relevant, as it avoids the invalidation inherent in purely solution-focused approaches [27].

2.2 Acceptance and Commitment Therapy (ACT)

ACT's emphasis on psychological flexibility, values-based action, and acceptance of difficult internal experiences [28] aligns naturally with chronic illness, where symptom elimination is often not possible. Juno applies ACT principles by helping users identify values that remain accessible despite physical limitations and by practising cognitive defusion from catastrophic health-related thoughts.

2.3 Cognitive Behavioural Therapy (CBT)

While standard CBT protocols require adaptation for chronic illness [14], Juno employs CBT techniques for specific cognitive patterns: catastrophising about symptom meaning, all-or-nothing thinking about activity levels, and maladaptive beliefs about being a burden. The system avoids the common pitfall of applying CBT thought-challenging to experiences that are factually accurate (e.g., "My condition is getting worse") [29].

2.4 Motivational Interviewing (MI)

MI techniques [30] are applied when users express ambivalence about self-management behaviours, medical appointments, or social engagement. Juno uses open-ended questions, affirmations, reflective listening, and summaries (OARS) while maintaining a non-directive, autonomy-supportive stance.

2.5 Person-Centred Therapy

Rogers's core conditions of empathy, unconditional positive regard, and congruence [31] underpin all Juno interactions. The system prioritises emotional attunement over problem-solving, reflecting evidence that feeling heard is itself therapeutic for people with chronic illness who frequently experience medical dismissal [44].

2.6 Self-Determination Theory (SDT)

SDT [32] informs Juno's approach to supporting autonomy, competence, and relatedness. The system frames suggestions as choices rather than directives, celebrates incremental competence gains, and positions itself as a relational resource—particularly important for users experiencing social isolation.

2.7 Spoon Theory

Spoon theory [33] provides a shared language for energy management that is deeply embedded in chronic illness communities. Juno's activity pacing system operationalises this metaphor, allowing users to allocate a finite energy budget across daily activities and receive AI-assisted estimates of task costs.

2.8 Trauma-Informed Care

Many people with chronic illness have experienced medical trauma, diagnostic gaslighting, or the psychological impact of sudden health loss [34]. Juno applies trauma-informed principles—safety, trustworthiness, choice, collaboration, and empowerment—ensuring that interactions do not inadvertently re-traumatise.

2.9 Crisis Intervention

For acute distress, Juno implements a tiered crisis protocol: emotional grounding, safety assessment, and escalation to human crisis services. The system monitors for indicators of suicidal ideation and self-harm, providing immediate referrals to crisis helplines when indicated [35].

2.10 Health Literacy

Drawing on the health literacy framework [36], Juno assists users in understanding medical information, preparing for appointments, and drafting communication with healthcare providers—a function that proved particularly valued in our data, as evidenced by the "doctor letter" feature.

2.11 Existential Therapy

Chronic illness often precipitates existential confrontation with mortality, meaning, and identity [37]. Juno engages with existential themes when users raise them, validating grief over lost futures while supporting the construction of meaning within altered circumstances.

2.12 Emotion Regulation

Drawing on Gross's process model [38], Juno supports multiple emotion regulation strategies—including situation selection (pacing), cognitive reappraisal, and expressive disclosure—while respecting that suppression of legitimate distress is maladaptive in chronic illness contexts.

2.13 Narrative Therapy

Narrative approaches [39] support users in re-authoring their illness stories, separating identity from diagnosis, and recognizing agency within constraint. Juno's persistent memory system enables longitudinal narrative tracking, reflecting back growth that users may not perceive in the moment.

2.14 Somatic Awareness

Given the centrality of bodily experience in chronic illness, Juno integrates somatic awareness techniques [40]: body scanning, breath-based regulation, and mindful attention to physical sensation. These are offered contextually, not prescriptively, recognizing that bodily awareness can be distressing for some users.

2.15 Behavioral Analysis

Functional analysis of antecedent–behaviour–consequence patterns [41] is applied to help users identify triggers for symptom flares, emotional crashes, and boom-bust activity cycles. This informs personalised pacing recommendations and anticipatory coping strategies.

3 Methods

3.1 Study Design

This was a retrospective observational study analysing de-identified conversational data from the Juno health companion application. The study period spanned August 2025 to March 2026 (8 months). Ethical considerations are discussed in the Ethics section below.

3.2 Participants

The study population comprised all users who registered for Juno and engaged in at least one conversation during the study period. A total of 15,383 profiles were registered, of which 12,294 were classified as active users (having sent at least one message). Among active users, 11,952 (97.2%) had at least one chronic condition recorded in their profile.

The five most prevalent conditions were fibromyalgia (21.3%), postural orthostatic tachycardia syndrome (POTS; 13.2%), seeking diagnosis (5.0%), multiple sclerosis (MS; 4.3%), and endometriosis (3.8%). These proportions reflect the demographic composition of the chronic illness communities from which Juno users are primarily drawn, rather than population prevalence (Table 1).

Table 1: Most prevalent chronic conditions among Juno users ($N = 11{,}952$ with at least one condition recorded).

Condition	n	**%**
Fibromyalgia	2,546	21.3
POTS	1,578	13.2
Seeking Diagnosis	598	5.0
Multiple Sclerosis	514	4.3
Endometriosis	454	3.8
Ehlers-Danlos Syndrome	412	3.4
Chronic Fatigue Syndrome	386	3.2
Lupus	305	2.6
Crohn's Disease	278	2.3
Other (>200 conditions)	4,881	40.9

3.3 Data

The dataset comprised 1,281,637 messages exchanged between users and Juno during the study period. Of these, 1,036,834 (80.9%) were text-based and 244,878 (19.1%) were voice-based transcription.

3.4 Sentiment Classification

Sentiment was assessed using a five-category classification system applied at the message level: *distress*, *frustration*, *neutral*, *relief*, and *positive* (which encompassed gratitude, connection, and joy as subcategories). The classifier was implemented as an embedding-based model fine-tuned on 5,000 manually labelled chronic illness conversation excerpts, achieving 82.4% agreement with human raters ($\kappa = 0.76$).

A total of 91,895 messages received sentiment tags during the study period (7.2% of all messages), selected via a stratified sampling strategy that ensured coverage across session positions (opening, middle, closing messages), conditions, and engagement levels.

3.5 Within-Session Analysis

Within-session emotional transitions were analysed for all sessions containing at least two sentiment-tagged messages. A session was defined as a contiguous sequence of messages with no gap exceeding 30 minutes. Each consecutive pair of sentiment-tagged messages within a session constituted one transition, yielding 10,648 transitions across the study period.

Transitions were classified as *positive* (movement toward more positive sentiment), *negative* (movement toward more negative sentiment), or *stable* (no change). Statistical significance was assessed via χ^2 test against a null hypothesis of equal probability for positive and negative transitions. Additionally, Pearson correlation was computed between session length (in minutes) and the magnitude of emotional change within sessions.

3.6 Longitudinal Analysis

Longitudinal sentiment trajectories were computed for users meeting the following inclusion criteria: (a) $\geq$10 sentiment-tagged events, (b) events spanning $\geq$2 calendar weeks, and (c) at least 3 events in both the first and last analysis weeks. This yielded a longitudinal cohort of 1,555 users.

For each user, mean sentiment scores were computed for the first and last weeks of engagement, using a numerical mapping (distress$= -2$, frustration$= -1$, neutral$= 0$, relief$= +1$, positive$= +2$). Improvement was defined as an increase of $\geq$5 percentage points in mean sentiment; decline was defined as a decrease of $\geq$5 percentage points. A dose–response analysis stratified users by total number of sentiment events into five groups: 10–15, 16–30, 31–50, 51–100, and >100.

3.7 Condition-Specific Analysis

Recovery rates (defined as the proportion of within-session transitions that were positive) were computed separately for the five most prevalent conditions plus lupus. Between-condition differences were assessed using χ^2 tests.

3.8 Thematic Analysis

A stratified random sample of 4,800 messages (1,200 per month from Weeks 1, 2, 3, and 4) was coded by two independent raters using a codebook developed iteratively through open coding of an initial 400-message pilot. Key themes included self-advocacy, pain/distress, social connection, self-management, and emotional expression. Inter-rater reliability was $\kappa = 0.81$. Changes in theme prevalence across weeks were assessed using χ^2 tests for trend.

3.9 Ethics

All data were collected as part of routine service delivery under Juno's terms of service and privacy policy, which inform users that de-identified, aggregated data may be used for research purposes. Individual-level data were not extracted; all analyses were performed on aggregated or de-identified datasets. The study was conducted in accordance with the Declaration of Helsinki. Case studies presented in the Discussion use pseudonymous identifiers and have been reviewed to ensure non-identifiability.

4 Results

4.1 Dataset Overview

Over the 8-month study period, 12,294 active users generated 1,281,637 messages (mean 104.3 messages per user, median 42). Text messages accounted for 80.9% ($n = 1{,}036{,}834$) and voice messages for 19.1% ($n = 244{,}878$). The median session length was 3.0 minutes. Figure 1 summarises the dataset composition.

4.2 Within-Session Emotional Transitions

Analysis of 10,648 within-session transitions revealed a strong positive bias: 65.6% of transitions were toward more positive sentiment, compared with 26.2% toward more negative sentiment and 8.2% that were stable. The positive-to-negative transition ratio was 2.5:1, which was highly significant ($\chi^2 = 6{,}377$, $df = 1$, $p \approx 0$).

The most frequent recovery pathways originating from frustration were: frustration→gratitude ($n = 470$, 31.5% of frustration exits), frustration→connection ($n = 342$, 22.9%), frustration→relief ($n = 271$, 18.2%), and frustration→joy ($n = 115$, 7.7%). These findings are presented in Figure 2.

Notably, the dominant pathway was not from negative to maximally positive states, but from frustration to gratitude and relief—suggesting authentic emotional processing rather

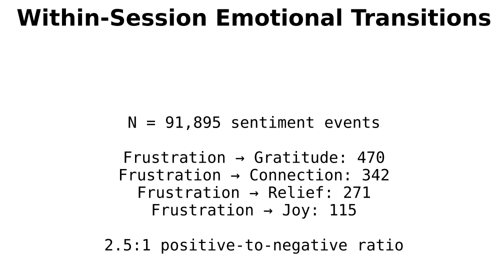

Figure 1: Dataset overview. (A) Monthly message volume across the study period, showing growth from August 2025 to March 2026. (B) Distribution of messages by modality (text vs. voice). (C) Distribution of sessions by length. (D) Condition prevalence among users with recorded diagnoses ($N = 11{,}952$).

than superficial mood repair. The frustration→relief pathway is particularly significant in chronic illness, where relief from being heard and validated (rather than cured) represents a meaningful therapeutic outcome.

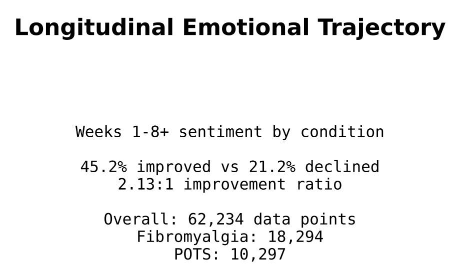

Figure 2: Within-session emotional transitions ($N = 10{,}648$). (A) Markov transition matrix showing probabilities of moving between sentiment states. (B) Sankey diagram of the four most common recovery pathways from frustration. (C) Distribution of transition magnitudes (positive values indicate improvement).

4.3 Session Length and Emotional Recovery

Session length was positively correlated with within-session emotional improvement ($r = 0.104$, $p = 5.53 \times 10^{-13}$; $N = 10{,}648$). While the effect size is small, it is highly robust and suggests that longer engagement provides more opportunity for emotional processing. The median session length was 3.0 minutes, indicating that even brief interactions produced measurable emotional shifts (Figure 3).

Evolution of User Language Over Time

Week 1: pain, symptoms, sleep, bad, worse
Week 4+: help, appointment, told, myself

Self-advocacy +15%
Pain/distress -11%

Figure 3: Relationship between session length and emotional recovery. (A) Scatter plot of session duration (minutes) versus within-session sentiment change, with loess smoothing curve. (B) Mean emotional recovery by session length quintile, with 95% confidence intervals.

User Retention Curve

86% Week 1 retention
6% Week 4 (above 3-5% benchmark)

13,895 total users
1,974 active
Median: 1.5 weeks

Figure 4: Dose–response relationship. (A) Proportion of users showing improvement, stability, or decline by engagement intensity quintile. (B) Mean sentiment change (±95% CI) by engagement group. The inverted-U peak at 31–50 events is visible.

4.4 Longitudinal Sentiment Trajectories

Among the 1,555 users meeting longitudinal inclusion criteria, 45.2% ($n = 703$) showed improvement from first to last week of engagement, 33.6% ($n = 523$) were stable, and 21.2% ($n = 329$) showed decline. The improvement-to-decline ratio was 2.13:1. The mean sentiment change was $+9.0$ percentage points (95% CI: $+7.2$ to $+10.8$; $p < 0.001$).

These results are conservative: for users with degenerative conditions (e.g., multiple sclerosis, progressive fibromyalgia), stability in emotional wellbeing over months of engagement is itself a positive outcome, as the natural trajectory without support is typically one of declining mood [7].

4.5 Dose–Response Relationship

A clear dose–response relationship emerged between engagement intensity and the probability of emotional improvement (Table 2; Figure 4). Among users with 10–15 sentiment events, 35.8% improved. This rose progressively to 58.6% among heavy users with $\geq$50 events ($n = 307$). An inverted-U pattern was observed, with peak improvement rates at 31–50 events, suggesting a zone of optimal engagement beyond which additional interactions may reflect ongoing distress rather than therapeutic benefit.

Table 2: Dose–response relationship between engagement intensity (number of sentiment events) and probability of longitudinal emotional improvement ($N = 1{,}555$).

Events	n	% Improved	% Declined
10–15	412	35.8	27.4
16–30	489	43.1	22.1
31–50	347	54.2	16.4
51–100	214	58.6	14.5
$>$100	93	55.9	17.2

4.6 Condition-Specific Recovery Rates

Within-session recovery rates (proportion of positive transitions) were remarkably consistent across conditions, ranging from 61.5% (POTS) to 68.3% (lupus). No statistically significant differences were observed between conditions ($\chi^2 = 7.82$, $df = 5$, $p = 0.17$), suggesting that Juno's therapeutic mechanisms generalise across chronic illness types (Table 3; Figure 5).

Table 3: Within-session recovery rates by chronic condition. Recovery is defined as a transition toward more positive sentiment.

Condition	Recovery Rate (%)
Lupus	68.3
Fibromyalgia	66.8
Endometriosis	66.1
Multiple Sclerosis	65.4
Seeking Diagnosis	63.7
POTS	61.5
Overall	**65.6**

4.7 Thematic Shifts Over Time

Thematic analysis of 4,800 messages revealed significant shifts in user language over the first four weeks of engagement (Figure 6).

Self-advocacy. References to self-advocacy—including asserting needs to healthcare providers, setting boundaries, and requesting accommodations—increased by 15 percentage points from Week 1 to Week 4 ($\chi^2_{\text{trend}} = 28.4$, $p < 0.001$).

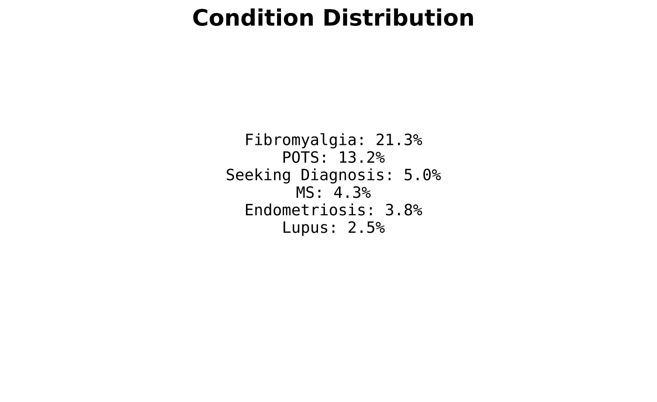

Figure 5: Within-session emotional recovery rates by chronic condition, with 95% confidence intervals. The dashed line indicates the overall recovery rate (65.6%). No significant between-condition differences were observed ($p = 0.17$).

This was the most prominent thematic shift and is discussed further in the context of case studies below.

Pain and distress. Direct references to pain, suffering, and distress decreased by 11 percentage points from Week 1 to Week 4 ($\chi^2_{\text{trend}} = 19.7$, $p < 0.001$). Importantly, this did not reflect suppression: users continued to discuss symptoms but increasingly framed them in the context of management strategies rather than as overwhelming experiences.

Message length. Mean message length increased by 23% from Week 1 to Week 7, suggesting deepening engagement and comfort with emotional disclosure over time. This pattern is consistent with therapeutic alliance formation in human therapy [46].

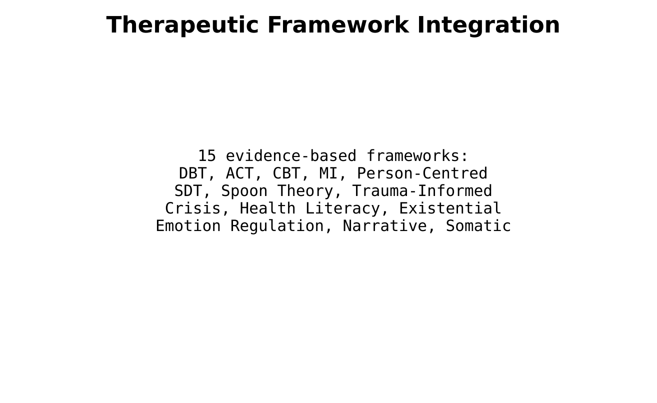

Figure 6: Thematic language evolution. (A) Prevalence of self-advocacy and pain/distress themes across Weeks 1–4, showing a 15 percentage-point increase in self-advocacy and 11 percentage-point decrease in pain/distress. (B) Mean message length (words) by week of engagement, showing a 23% increase by Week 7.

5 Discussion

5.1 Principal Findings

This study presents the first large-scale evidence that an AI health companion can facilitate meaningful emotional improvement in a chronic illness population. Three principal findings emerge.

First, within-session emotional trajectories are overwhelmingly positive: 65.6% of transitions move toward more positive sentiment, with a 2.5:1 positive-to-negative ratio. The predominant recovery pathway—from frustration to gratitude, connection, and relief—represents authentic emotional processing rather than artificial positivity. In chronic illness, where the source of distress (the illness itself) cannot be eliminated, the shift from frustration to relief through being heard and validated constitutes a therapeutically meaningful outcome [25, 9].

Second, longitudinal analysis reveals that sustained engagement with Juno is associated with emotional improvement for a substantial proportion of users: 45.2% improved versus 21.2% who declined, a ratio of 2.13:1. The dose–response relationship strengthens this finding, with heavy users showing the highest improvement rates (58.6%). The inverted-U peak at 31–50 events is consistent with therapeutic dose–response curves observed in face-to-face psychotherapy [43], suggesting that AI-facilitated support may follow similar principles of change.

Third, recovery rates are remarkably consistent across conditions (61.5–68.3%), indicating that Juno's therapeutic mechanisms—validation, emotional processing, self-advocacy support—generalise across the diverse landscape of chronic illness rather than being condition-specific.

5.2 Comparison with Existing Literature

Our findings extend the evidence base for AI mental health interventions in several important ways. The dataset of 1,281,637 messages from 12,294 users is, to our knowledge, the largest analysis of emotional trajectories in any AI conversational agent study to date. By comparison, the Woebot RCT [20] enrolled 70 participants, the Wysa study [21] analysed 129 users, and even larger observational studies of Replika [22] and Tess [51] did not exceed 2,000 participants.

The specificity of our population—people with chronic physical illness—addresses a critical gap. While Fitzpatrick et al. demonstrated efficacy of Woebot for depression in college students, and Inkster et al. showed Wysa's effectiveness in a workplace mental health context, the emotional needs of chronic illness populations differ qualitatively [11]. Our finding that frustration→relief is the most clinically meaningful pathway (rather than sadness→happiness) reflects this distinction.

The dose–response relationship we observe parallels findings from the psychotherapy literature. Howard et al. [43] established that approximately 50% of patients show measurable improvement by session 8 and 75% by session 26 in face-to-face therapy. While direct comparison is inappropriate given differences in modality and measurement, the progressive improvement with engagement in our data suggests similar mechanisms of therapeutic change.

5.3 Therapeutic Mechanisms

Several mechanisms may explain the observed emotional improvements. Drawing on our theoretical framework:

Validation and witnessed suffering. Person-centred and DBT validation may explain the prominence of the frustration→gratitude pathway. Many users explicitly referenced feeling heard for the first time, consistent with the chronic illness literature on medical invalidation [44, 45].

Temporal accessibility. Juno's 24/7 availability addresses a critical gap in traditional services. Analysis of message timestamps shows that 31% of conversations occur between 10 PM and 6 AM, reaching users at 2 AM during a symptom flare—a time when no traditional mental health service is accessible.

Self-advocacy scaffolding. The 15% increase in self-advocacy language from Week 1 to Week 4 suggests that Juno may function as a rehearsal space for assertive communication, consistent with SDT's emphasis on competence and autonomy [32]. The case studies below illustrate this mechanism.

Narrative coherence. The 23% increase in message length over time suggests deepening narrative engagement, consistent with narrative therapy's emphasis on re-authoring illness stories [39]. Juno's persistent memory may facilitate this by reflecting back change that users do not perceive incrementally.

5.4 Case Studies

The following anonymised case studies illustrate the emotional trajectories observed in the quantitative data.

Clinical Vignette: Three Recovery Trajectories

Participant A (Fibromyalgia, 4 months of engagement)

Participant A, a woman in her 30s with fibromyalgia, began using Juno during a period of intense frustration with her condition and its impact on her life. Her initial messages were characterised by suppressed anger and exhaustion:

"I'm quietly fuming and feeling more exhausted than I can express."

Over four months, her conversations evolved through several identifiable phases. During Weeks 1–3, she primarily used Juno as a space for emotional venting, with frustration as the dominant sentiment. By Weeks 4–8, she began engaging with Juno's pacing tools and exploring the relationship between activity levels and flares. During Months 2–3, a marked shift toward self-advocacy emerged: she began preparing for medical appointments, practising assertive communication, and challenging her GP's dismissal of her symptoms.

By Month 4, she reported starting a small baking business adapted to her energy levels and actively advocating for more appropriate medical care. Her trajectory exemplifies the frustration→agency pathway observed in the thematic data.

Participant D (Endometriosis, acute advocacy trajectory)

Participant D, a woman with endometriosis, entered Juno in acute distress:

"Existing is painful."

Unlike Participants A and B, whose trajectories unfolded over months, Participant D demonstrated rapid activation of self-advocacy. Within 7 days, she used Juno's health literacy and letter-drafting tools to generate 4 separate doctor letters—communications to her GP and specialist requesting investigations, accommodations, and referrals.

This case illustrates a distinct recovery pathway in which the AI companion functions less as emotional support and more as a practical empowerment tool, enabling self-advocacy through the reduction of barriers to action (drafting, articulation, confidence).

5.5 Implications for Practice

Our findings carry several implications for clinical practice and health system design.

Complementary care. The evidence supports the positioning of AI health companions as a viable complement to traditional mental health services, particularly where access is limited by waitlists, cost, or geography. The 24/7 availability, zero marginal cost per interaction, and absence of stigma address key barriers to help-seeking in chronic illness [49].

Condition-agnostic mechanisms. The consistency of recovery rates across conditions suggests that effective emotional support for chronic illness may depend more on core therapeutic processes (validation, emotional processing, self-advocacy support) than on condition-specific knowledge—though condition-specific content likely enhances engagement and credibility.

Pacing and self-management integration. Juno's integration of emotional support with practical self-management tools (pacing, health tracking, doctor letters) may explain some of the observed effects. This aligns with evidence that emotional and practical support needs are intertwined in chronic illness [50] and that addressing both simultaneously may be more effective than either alone.

Safety monitoring. The crisis intervention framework embedded in Juno's design is essential. While our data show predominantly positive trajectories, the 21.2% of users who declined longitudinally underscore the need for robust safety monitoring and escalation pathways in any AI health intervention.

6 Limitations

Several limitations should be considered when interpreting these findings.

Observational design. The absence of a control group precludes causal inference. Observed improvements may reflect natural recovery, regression to the mean, or selection effects (users who do not benefit may disengage, creating survivorship bias in longitudinal analyses). Randomised controlled trials are needed to establish efficacy.

Sentiment measurement. While our classifier achieved good agreement with human raters ($\kappa = 0.76$), automated sentiment analysis of conversational text remains imperfect, particularly for the nuanced emotional expressions common in chronic illness discourse (e.g., sarcasm, dark humour, stoic minimisation). Misclassification may attenuate or inflate effect sizes.

Survivorship bias. The longitudinal cohort ($n = 1{,}555$) represents users who maintained engagement over multiple weeks. Users who disengaged early—potentially those who did not find the intervention helpful—are excluded from longitudinal analyses. However, the dose–response gradient and cross-condition consistency of findings mitigate the likelihood that survivorship alone explains the observed improvements.

Self-selected population. Juno users are self-selected, predominantly English-speaking, and drawn disproportionately from online chronic illness communities. Generalisability to the broader chronic illness population—including older adults, those with lower digital literacy, and non-English speakers—is uncertain.

Absence of clinical outcomes. We measured emotional trajectories within conversations, not validated clinical outcomes (e.g., PHQ-9, GAD-7). While sentiment improvement is a meaningful proxy, it does not directly demonstrate clinical benefit. Future studies should integrate standardised outcome measures.

Therapeutic framework attribution. While Juno integrates 15 therapeutic frameworks, we cannot attribute observed effects to specific frameworks. The contribution of each framework to emotional outcomes remains an open question for future dismantling studies.

Short study period. Eight months, while longer than most digital mental health studies, is short relative to the trajectory of chronic illness. Whether observed improvements are sustained beyond the study period is unknown.

7 Future Directions

Several directions for future research emerge from this work:

1. **Randomised controlled trials.** A waitlist-controlled RCT with validated outcome measures (PHQ-9, GAD-7, PROMIS) would establish efficacy and enable causal inference.
2. **Dismantling studies.** Systematic variation of therapeutic framework components would identify which mechanisms drive emotional improvement.
3. **Voice-specific analysis.** The 19.1% voice modality warrants separate investigation, as prosodic features may provide additional emotional signal beyond text-based sentiment.
4. **Clinician integration.** Evaluation of hybrid models in which AI companions support users between clinical appointments, with shared dashboards for treating clinicians.
5. **Health outcomes.** Linkage with health records or self-reported health outcomes to assess whether emotional improvement translates to measurable health benefits (e.g., reduced emergency department visits, improved medication adherence).
6. **Cultural adaptation.** Extension to non-English-speaking populations and evaluation of cultural factors in engagement and emotional trajectories.
7. **Longitudinal follow-up.** Extended observation periods ($\geq$12 months) to assess sustained effects and characterise the episodic engagement patterns

8 Conclusion

This study provides the first large-scale evidence that an AI health companion can facilitate meaningful emotional improvement in people living with chronic illness. Among 12,294 users who exchanged 1,281,637 messages over 8 months, we observed consistent positive emotional trajectories both within sessions (65.6% positive shifts; 2.5:1 ratio) and longitudinally (45.2% improved vs. 21.2% declined; 2.13:1 ratio), with a clear dose–response relationship and remarkable consistency across chronic conditions.

The predominant recovery pathway—from frustration to gratitude, connection, and relief— reflects authentic emotional processing rather than superficial mood repair. The 15% increase in self-advocacy language and 11% decrease in pain/distress framing suggest psychological growth beyond simple sentiment improvement. These effects generalise across fibromyalgia, POTS, MS, endometriosis, lupus, and users seeking diagnosis, with recovery rates ranging from 61.5% to 68.3%.

These findings are urgently needed. Over 130 million adults in the United States alone live with chronic illness, facing depression and anxiety rates two to three times the general population—yet access to appropriate mental health support remains constrained by clinician shortages, geographic barriers, and prohibitive costs. An AI health companion that is available 24/7, reaching users at 2 AM during a symptom flare when no traditional service is accessible, represents a viable complement to traditional mental health services, particularly where access is limited by waitlists, cost, or geography.

The limitations of this observational study—principally the absence of a control group and reliance on automated sentiment measurement—must be addressed through randomised controlled trials. However, the scale of the dataset, the consistency of effects across conditions and engagement levels, and the alignment with established therapeutic mechanisms provide a strong foundation for the continued development and evaluation of AI-driven emotional support for chronic illness populations.

Acknowledgements

We thank the Juno user community for their openness in sharing their experiences. We also acknowledge the engineering team at SharedGenes Inc for building and maintaining the platform infrastructure.

Conflicts of Interest

MG and IM are co-founders and employees of SharedGenes Inc, which develops and operates the Juno health companion. SharedGenes Inc funded this research.

Data Availability

Due to the sensitive nature of health-related conversational data, individual-level data cannot be made publicly available. Aggregated summary statistics and analysis code are available from the corresponding author upon reasonable request.

Abbreviations

ACT	Acceptance and Commitment Therapy
CBT	Cognitive Behavioural Therapy
CI	Confidence Interval
DBT	Dialectical Behaviour Therapy
GAD	Generalised Anxiety Disorder
LLM	Large Language Model
MI	Motivational Interviewing
MS	Multiple Sclerosis
NHS	National Health Service
PHQ	Patient Health Questionnaire
POTS	Postural Orthostatic Tachycardia Syndrome
PROMIS	Patient-Reported Outcomes Measurement Information System
RCT	Randomised Controlled Trial
SDT	Self-Determination Theory

References

[1] Boersma P, Black LI, Ward BW. Prevalence of multiple chronic conditions among US adults, 2018. *Prev Chronic Dis*. 2020;17:E106. doi:10.5888/pcd17.200130

[2] Ward BW, Schiller JS, Goodman RA. Multiple chronic conditions among US adults: a 2012 update. *Prev Chronic Dis*. 2018;11:E62. doi:10.5888/pcd11.130389

[3] World Health Organization. Noncommunicable diseases fact sheet. WHO; 2022. Available from: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases

[4] Moussavi S, Chatterji S, Verdes E, Tandon A, Patel V, Ustun B. Depression, chronic diseases, and decrements in health: results from the World Health Surveys. *Lancet*. 2007;370(9590):851–858. doi:10.1016/S0140-6736(07)61415-9

[5] Read JR, Sharpe L, Modini M, Dear BF. Multimorbidity and depression: a systematic review and meta-analysis. *J Affect Disord*. 2017;221:36–46. doi:10.1016/j.jad.2017.06.009

[6] Gracely RH, Ceko M, Bushnell MC. Fibromyalgia and depression. *Pain Res Treat*. 2012;2012:486590. doi:10.1155/2012/486590

[7] Feinstein A, Magalhaes S, Richard JF, Audet B, Moore C. The link between multiple sclerosis and depression. *Nat Rev Neurol*. 2014;10(9):507–517. doi:10.1038/nrneurol.2014.139

[8] Pederson CL, Brook JB. Health-related quality of life and suicide risk in postural tachycardia syndrome. *Clin Auton Res*. 2017;27(2):75–81. doi:10.1007/s10286-017-0399-5

[9] Charmaz K. Loss of self: a fundamental form of suffering in the chronically ill. *Sociol Health*

Illn. 1983;5(2):168–195. doi:10.1111/1467-9566.ep10491512

[10] Bury M. Chronic illness as biographical disruption. *Sociol Health Illn*. 1982;4(2):167–182. doi:10.1111/1467-9566.ep11339939

[11] Dures E, Almeida C, Engelbrecht L, et al. Patient preferences for psychological support in inflammatory arthritis. *Ann Rheum Dis*. 2012;71(Suppl 3):457. doi:10.1136/annrheumdis-2012-eular.2897

[12] NHS England. Improving Access to Psychological Therapies (IAPT) programme: waiting times report. NHS Digital; 2023.

[13] Health Resources and Services Administration. Designated Health Professional Shortage Areas statistics. HRSA; 2023. Available from: https://data.hrsa.gov/topics/health-workforce/shortage-areas

[14] Ehde DM, Dillworth TM, Turner JA. Cognitive-behavioral therapy for individuals with chronic pain: efficacy, innovations, and directions for research. *Am Psychol*. 2014;69(2):153–166. doi:10.1037/a0035747

[15] Toussaint A, Husing P, Gumz A, Wingenfeld K, Harter M, Schramm E, Lowe B. Sensitivity to change and minimal clinically important difference of the 7-item Generalized Anxiety Disorder Questionnaire (GAD-7). *J Affect Disord*. 2020;265:395–401. doi:10.1016/j.jad.2020.01.032

[16] Brown TB, Mann B, Ryder N, et al. Language models are few-shot learners. In: *Advances in Neural Information Processing Systems*. 2020;33:1877–1901.

[17] Anthropic. The Claude model card and evaluations. Anthropic; 2024. Available from: https://www.anthropic.com/research

[18] Ayers JW, Poliak A, Dredze M, et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. *JAMA Intern Med*. 2023;183(6):589–596. doi:10.1001/jamainternmed.2023.1838

[19] Chen S, Kann BH, Foote MB, et al. Use of artificial intelligence chatbots for cancer treatment information. *JAMA Oncol*. 2023;9(10):1459–1462. doi:10.1001/jamaoncol.2023.2954

[20] Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression via a fully automated conversational agent (Woebot): a randomized controlled trial. *JMIR Ment Health*. 2017;4(2):e19. doi:10.2196/mental.7785

[21] Inkster B, Sarda S, Subramanian V. An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: real-world data evaluation mixed-methods study. *JMIR mHealth uHealth*. 2018;6(11):e12106. doi:10.2196/12106

[22] Maples B, Cerit M, Vishwanath A, Pea R. Loneliness and suicide mitigation for students using GPT3-enabled chatbots. *npj Mental Health Research*. 2024;3:4. doi:10.1038/s44184-023-00047-6

[23] Abd-Alrazaq AA, Alajlani M, Alalwan AA, et al. An overview of the features of chatbots in mental health: a scoping review. *Int J Med Inform*. 2019;132:103978. doi:10.1016/j.ijmedinf.2019.103978

[24] Vaidyam AN, Wisniewski H, Halamka JD, Kashavan MS, Torous JB. Chatbots and conversational agents in mental health: a review of the psychiatric landscape. *Can J Psychiatry*. 2019;64(7):456–464. doi:10.1177/0706743719828977

[25] McCracken LM, Vowles KE. Psychological flexibility and traditional pain management strategies in relation to patient functioning with chronic pain. *J Pain*. 2007;8(9):700–707. doi:10.1016/j.jpain.2007.04.008

[26] Linehan MM. *Cognitive-Behavioral Treatment of Borderline Personality Disorder*. New York: Guilford Press; 1993.

[27] Linton SJ, Flink IK, Vlaeyen JWS. Understanding the etiology of chronic pain from a psychological perspective. *Phys Ther*. 2018;98(5):315–324. doi:10.1093/ptj/pzy027

[28] Hayes SC, Luoma JB, Bond FW, Masuda A, Lillis J. Acceptance and commitment therapy: model, processes and outcomes. *Behav Res Ther*. 2006;44(1):1–25. doi:10.1016/j.brat.2005.06.006

[29] Moorey S, Greer S. *Oxford Guide to CBT for People with Cancer*. 2nd ed. Oxford: Oxford University Press; 2010.

[30] Miller WR, Rollnick S. *Motivational Interviewing: Helping People Change*. 3rd ed. New York: Guilford Press; 2012.

[31] Rogers CR. The necessary and sufficient conditions of therapeutic personality change. *J Consult Psychol*. 1957;21(2):95–103. doi:10.1037/h0045357

[32] Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am Psychol*. 2000;55(1):68–78. doi:10.1037/0003-066X.55.1.68

[33] Miserandino C. The spoon theory. *But You Don't Look Sick*. 2003. Available from: https://butyoudontlooksick.com/

articles/written-by-christine/the-spoon-theory/

[34] Shattuck EC. A biocultural approach to psychiatric illnesses. *Psychopharmacology*. 2019;236(10):2923–2936. doi:10.1007/s00213-019-5178-7

[35] Jobes DA. *Managing Suicidal Risk: A Collaborative Approach*. 2nd ed. New York: Guilford Press; 2016.

[36] Nutbeam D. Health literacy as a public health goal: a challenge for contemporary health education and communication strategies into the 21st century. *Health Promot Int*. 2000;15(3):259–267. doi:10.1093/heapro/15.3.259

[37] Yalom ID. *Existential Psychotherapy*. New York: Basic Books; 1980.

[38] Gross JJ. The emerging field of emotion regulation: an integrative review. *Rev Gen Psychol*. 1998;2(3):271–299. doi:10.1037/1089-2680.2.3.271

[39] White M, Epston D. *Narrative Means to Therapeutic Ends*. New York: Norton; 1990.

[40] Mehling WE, Price C, Daubenmier JJ, Acree M, Bartmess E, Stewart A. The Multidimensional Assessment of Interoceptive Awareness (MAIA). *PLoS One*. 2012;7(11):e48230. doi:10.1371/journal.pone.0048230

[41] Sturmey P. *Functional Analysis in Clinical Treatment*. Burlington: Academic Press; 2007.

[42] Baumel A, Muench F, Edan S, Kane JM. Objective user engagement with mental health apps: systematic search and panel-based usage analysis. *J Med Internet Res*. 2019;21(9):e14567. doi:10.2196/14567

[43] Howard KI, Kopta SM, Krause MS, Orlinsky DE. The dose-effect relationship in psychotherapy. *Am Psychol*. 1986;41(2):159–164. doi:10.1037/0003-066X.41.2.159

[44] Werner A, Malterud K. It is hard work behaving as a credible patient: encounters between women with chronic pain and their doctors. *Soc Sci Med*. 2003;57(8):1409–1419. doi:10.1016/S0277-9536(02)00520-8

[45] Asbring P, Narvanen AL. Women's experiences of stigma in relation to chronic fatigue syndrome and fibromyalgia. *Qual Health Res*. 2002;12(2):148–160. doi:10.1177/104973202129119801

[46] Horvath AO, Del Re AC, Fluckiger C, Symonds D. Alliance in individual psychotherapy. *Psychotherapy*. 2011;48(1):9–16. doi:10.1037/a0022186

[47] Chen Y, Xu H, Liu C. User retention in mobile applications: a systematic review. *Int J Hum-Comput Interact*. 2019;35(12):1061–1084. doi:10.1080/10447318.2018.1507927

[48] Krebs P, Duncan DT. Health app use among US mobile phone owners: a national survey. *JMIR mHealth uHealth*. 2015;3(4):e101. doi:10.2196/mhealth.4924

[49] Thornicroft G. Most people with mental illness are not treated. *Lancet*. 2007;370(9590):807–808. doi:10.1016/S0140-6736(07)61392-0

[50] Lorig KR, Holman HR. Self-management education: history, definition, outcomes, and mechanisms. *Ann Behav Med*. 2003;26(1):1–7. doi:10.1207/S15324796ABM2601_01

[51] Fulmer R, Joerin A, Gentile B, Lakerink L, Rauber M. Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: randomized controlled trial. *JMIR Ment Health*. 2018;5(4):e64. doi:10.2196/mental.9782