Exploratory Study for Reduction of Anxiety and Depression Through Dual Mental Health Treatment: Digital-Based Cognitive Behavioral Therapy and National Board Health and Wellness Coaches (NBHWC) For People With Disabilities

NCT: NCT06477913

Date: July 1, 2026

Objective: This pilot study evaluated Rauha, a novel digital therapeutic created by Toivoa Inc, that integrates cognitive behavioral therapy (CBT)–based modules with live weekly sessions led by a National Board-Certified Health and Wellness Coach (NBC-HWC), delivering structured, smartphone-based psychoeducation and interactive therapeutic exercises combined with personalized mental health coaching to support behavior change.

Design: This study used a single-arm, within-participants design to measure the clinical impact of Rauha on symptoms of anxiety and depression, as measured by the Hamilton Anxiety Rating Scale (HAM-A) and Hamilton Depression Rating Scale (HAM-D), respectively. Patients were assessed with the HAM-A and HAM-D at baseline (preintervention), immediately after completing treatment (post intervention), and 4 weeks after completing treatment (follow-up). Participation from baseline to follow-up lasted approximately 12 weeks.

Methods: Interested adults completed a pre-consent screening via the GAD-7, PHQ-9, and self-reported disability status. Eligible individuals underwent full informed consent and baseline assessment using the clinician-administered Hamilton Anxiety Rating Scale (HAM-A) and Hamilton Depression Rating Scale (HAM-D) before completing a virtual orientation to onboard onto the Rauha digital therapeutic (DTx) smartphone application. The intervention consisted of an 8-week program combining self-guided, CBT-based smartphone modules (~60 minutes/week) with weekly 30-minute virtual sessions led by a National Board Certified Health and Wellness Coach (NBC-HWC). Coaching sessions followed a standardized, semistructured framework focusing on mood check-ins, skill reinforcement, and goal setting, preceded by mandatory safety monitoring using the Columbia-Suicide Severity Rating Scale (C-SSRS) with an established clinical escalation

protocol. To maintain coaching fidelity, coaches completed standardized training and ongoing supervision by a licensed clinician. Follow-up clinical assessments (HAM-A and HAM-D) were administered post-treatment (Week 8) and at a 4-week follow-up (Week 12); during this post-intervention period, participants maintained independent access to the mobile application. To minimize assessment variability, independent, licensed clinicians administered all assessments, with each participant evaluated by the same rater across all time points.

Statistical Analysis Plan: The primary retention rate was calculated as the proportion of participants completing all 8-week smartphone modules and coaching sessions. Main clinical efficacy analyses utilized an intent-to-treat approach including all participants who initiated treatment. Continuous variables are summarized using means and standard deviations (SDs), and categorical data using percentages. Mixed-effects modeling (MEM) with a compound symmetry variance-covariance structure was applied to analyze repeated-measures symptom changes across discrete time points (baseline, post-treatment, and follow-up), utilizing linear contrasts to evaluate intervention and durability effects. Within-person effect sizes were quantified using partial eta squared. Clinical Significance Reference (CSR) proportions, symptom percent changes (ranging from 10% to 90%), and associated 95% confidence intervals were calculated using Clopper-Pearson exact formulas. Participant-reported survey responses for acceptability, usability, and user experience were evaluated descriptively via item- and scale-level Likert-scale endorsement percentages. App engagement was explored by calculating the average cumulative lesson replay rate per participant via passive platform analytics. Pre-planned power analyses determined that the sample size provided approximately 79% to 87% power ($\alpha = .05$, two-tailed) to detect historical target score reductions on the HAM-D and HAM-A. Given the exploratory proof-of-concept nature of this pilot study, no corrections for multiple comparisons were applied, and significance was set a priori at $p \leq .05$.