

Article

AI-generated and human-optimized personalized music interventions for anxiety relief in oncology nursing care: A novel approach combining technology and medicine

Haode Jian

Yonsei University, Seoul 03722, South Korea; gdstudyveryqiang@naver.com

CITATION

Jian H. AI-generated and human-optimized personalized music interventions for anxiety relief in oncology nursing care: A novel approach combining technology and medicine. *Psycho-Oncologie*. 2025; x(x): 4592.
<https://doi.org/10.18282/po4592>

ARTICLE INFO

Received: 17 June 2025
Revised: 28 November 2025
Accepted: 17 December 2025
Available online: 30 December 2025

COPYRIGHT



Copyright © 2026 by author(s).
Psycho-Oncologie is published by PiscoMed Publishing Pte Ltd. This work is licensed under the Creative Commons Attribution (CC BY) license.
<https://creativecommons.org/licenses/by/4.0/>

Abstract: Music has been used for centuries to modulate human emotional states and reduce psychological stress and therefore holds great potential within the emerging field of psychosocial oncology care. In this paper we explore whether a treatment efficacy profile can be ascribed to a hybrid model of personalized music intervention combining AI-generated compositions with manual refinement provided by trained experts in order to reduce anxiety in cancer patients. The results indicate that, when compared to standardised music interventions, when compared to a defined standard music intervention control in this trial, the hybrid intervention demonstrated statistically significant reductions across a number of parameters including anxiety scores, physiological parameters, and general patient satisfaction. The mechanism of action seems to be mediated via a synergistic framework which engages both affective properties and neural regulatory processes, allowing for a multilayered and adaptive engagement with the treatment. This approach not only demonstrates the potential value of music as a non-pharmacological treatment, but also provides a scalable patient-centric intervention that extends the current repertoire of supportive care in oncology.

Keywords: AI-generated music; cancer-related anxiety; personalized intervention; randomized controlled trial

1. Introduction

Cancer is a highly stressful and life-threatening disease that inflicts not only severe physical damage but also prolonged psychological distress. Among various psychosocial factors, anxiety stands out as one of the most prevalent yet frequently underrecognized affective disorders in cancer patients [1,2]. Empirical studies indicate that over 40% of oncology patients experience varying degrees of anxiety throughout different phases of their disease trajectory—from initial diagnosis through treatment and follow-up care [3]. This persistent anxiety significantly compromises treatment adherence, impairs immune function, and substantially diminishes patients' quality of life. Despite the clinical effectiveness of conventional anxiety interventions (such as pharmacotherapy and psychological counseling), their application in clinical oncology is often limited by side effects, shortage of institutional resources and low patient acceptance [4,5]. In contrast, owing to its non-invasive regulation of emotions, sleep and neuroendocrine activity, music intervention has received increasing attention in clinical practice, while most of the current applications still depend on fixed music play lists, which are usually insensitive to the affective states and interests of individuals, resulting in limited emotional resonance and treatment effects [6,7]. Recent advancements in AI,

particularly deep learning models for generative music, have opened new avenues for creating highly personalized auditory stimuli. These models can be conditioned on specific emotional descriptors or physiological states, offering a level of customization previously unattainable with standard music libraries [8,9]. Concurrently, a growing body of evidence continues to elucidate the psychophysiological mechanisms of music, demonstrating its modulatory effects on the autonomic nervous system and the hypothalamic-pituitary-adrenal axis, which are often dysregulated in cancer-related anxiety [10]. This study builds upon this converging evidence by integrating state-of-the-art AI music generation into a clinically optimized framework.

With the rapid development of artificial intelligence technology, we explored the possibility of using machine learning methods to generate personalized music compositions in clinical practice [11,12]. Specifically, music generation models based on artificial intelligence were used to compose emotionally consistent and acoustically suitable pieces according to the affective cues and auditory preferences of user [13]. Further, the compositions were polished by experienced musicians to achieve better clinical application potential and patient acceptability. In this study, we proposed to explore the clinical value of a personalized music intervention model based on artificial intelligence music composition and human-in-the-loop optimization in alleviating anxiety in cancer patients. Specifically, we studied its clinical application value in multiple psychological indicators, physiological parameters and subjective experience. Therefore, this study aimed to directly test the primary hypothesis that a personalized music intervention model, combining AI-generated compositions based on individual patient profiles with subsequent manual refinement by music therapists, is superior to a standardized music intervention in reducing anxiety levels (as measured by psychological scales and physiological biomarkers) among cancer patients.

2. Materials and methods

2.1. Participants

2.1.1. Study population

In this study, we recruited 160 cancer patients with clinically evident anxiety symptoms (SAS score ≥ 50) admitted to the oncology department of a tertiary hospital from January 2023 to March 2025. These patients were distributed between the ages of 18 and 75 years, including 82 males and 78 females. All patients were pathologically diagnosed with malignant tumors, including breast cancer ($n = 42$), lung cancer ($n = 38$), colorectal cancer ($n = 32$), and other cancers ($n = 48$). The inclusion standards were KPS score ≥ 60 and an estimated survival time of more than 6 months. Patients with severe hearing impairment or intellectual impairment were excluded. This study was approved by the Institutional Review Board (Approval No. 2022-085). All patients provided a written informed consent form before entering the study.

2.1.2. AI model specifications and training

The AI music generation leveraged two primary models. For MusicLM (a

500M parameter transformer-based model), we utilized a pre-trained checkpoint which was subsequently fine-tuned on our curated corpus of 10,000 clinically validated relaxation music excerpts. The training data was sourced from proprietary therapeutic libraries and the Free Music Archive (FMA), annotated with valence-arousal tags. Key fine-tuning hyperparameters included a learning rate of $1e-4$, batch size of 16, and training for 50,000 steps. For OpenAI's Jukebox, we employed the pre-trained 5B parameter model (a hierarchical VQ-VAE architecture) and performed few-shot inference conditioned on the patient's emotional vector A and text prompts derived from preference questionnaires (e.g., "soothing piano melody"). This approach combined the robust prior knowledge of large-scale models with targeted conditioning for clinical applicability.

2.1.2. Inclusion and exclusion criteria

Inclusion criteria were as follows: (1) Histologically confirmed diagnosis of malignancy; (2) SAS (Self-Rating Anxiety Scale) score ≥ 50 ; (3) Age between 18 and 75 years; (4) No severe auditory or cognitive impairment; (5) Voluntary participation with signed informed consent. Exclusion criteria included: (1) History of psychiatric disorders or long-term use of anxiolytic medications; (2) Coexisting severe cardiac, pulmonary, hepatic, or renal dysfunction (e.g., NYHA Class III–IV heart failure); (3) Receipt of other music-based or psychological interventions within the past month; (4) Poor compliance with music therapy (e.g., refusal to wear headphones); (5) Pregnancy or lactation. Elimination criteria: Participants who withdrew due to clinical deterioration or were lost to follow-up during the intervention period were excluded. Ultimately, 8 cases were removed, yielding a final valid sample size of 152.

2.1.3. Sample size calculation and group allocation

Sample size was calculated by PASS 15.0 software. Assuming the effect size of anxiety score reduction was moderate based on previous reports [12], and $\alpha = 0.05$ (two-sided), $1 - \beta = 0.80$, the required sample size for each group was 72. Considering the expected dropout rate of 20%, 160 patients were included in the study. Stratified randomization was conducted according to the cancer type (breast/lung/colorectal/other) and SAS score [50–59] vs. ≥ 60 . Random sequences were generated by SPSS 26.0. The experimental group (AI-created music with manual optimization, $n = 80$) and control group (standard music intervention, $n = 80$) were enrolled. There were no significant differences in age, sex, stage of cancer, and initial SAS scores between the two groups ($P > 0.05$). The intervention lasted for four weeks. The number of sessions was three times a week. Each session lasted 30 min and was implemented within the 24 h after chemotherapy.

2.2. Intervention methods

2.2.1. AI-based music generation: Technical specifications

This research utilized Google's MusicLM and OpenAI Jukebox as the underlying AI models for generating the music. Both models use deep Transformer based methods to generate emotionally congruent music for individual patient profiles. Parameters for the input were labels of emotion, obtained from a blending

of SAS (Self-Rating Anxiety Scale) scores and emotional content of speech, as well as user preference data, collected from pre-intervention questionnaires on rhythm sensitivity and instrumental preferences. Specifically, the target emotional state vector A was constructed as follows: The SAS score was normalized and mapped to a valence-arousal coordinate within a two-dimensional emotional model, prioritizing low arousal and positive valence for relaxation. The emotional content of speech, obtained from a brief pre-intervention conversation about the patient's current mood, was analyzed using the OpenSMILE toolkit (version 3.0) to extract low-level acoustic features (e.g., pitch, intensity, spectral properties). These features were then classified into the same valence-arousal model using a pre-trained support vector machine (SVM) classifier. The final vector A was a weighted average of the SAS-derived coordinates and the speech-derived coordinates, giving slightly higher weight (0.6) to the SAS score due to its standardized nature. This quantitative process ensured a replicable input for the AI models. The corpus of clinically validated excerpts used for AI generation were over 10,000 clinically validated excerpts of music curated for relaxation and clinical use. All audio files were generated at 44.1 khz sampling rate at 16 bit depth, allowing for high-quality acoustic output suitable for clinical use. The emotional congruence of generated music acoustic stimuli with the patient's affect was measured via cosine similarity computations between vectors of multidimensional emotional embeddings of the music and target affective profiles. The similarity score informed the final selection and optimization stage prior to patient delivery, ensuring a high degree of affective matching:

$$\text{Similarity} = \frac{\sum_{i=1}^n (A_i \times B_i)}{\sqrt{\sum_{i=1}^n A_i^2} \times \sqrt{\sum_{i=1}^n B_i^2}}$$

here, vector A denotes the target emotional state of the patient, derived from SAS scores and real-time vocal affect analysis, while vector B represents the emotional feature vector of the generated music, extracted using OpenSMILE, an open-source toolkit for audio-based emotion recognition.

2.2.2. Manual optimization and customization workflow

Following AI-generated music creation, every piece was subject to a post-production human tuning process by two MT-BC board-certified music therapists for the following dual reasons: improvement of clinical sedation efficacy, and increase of patient music receptivity. This optimization was conducted using the digital audio workstation (DAW) Ableton Live 11 Suite. Specific technical procedures included: (1) Using the 'Warp' feature for precise tempo alignment to the 60-88 BPM range. (2) Applying EQ filters (e.g., high-pass filter below 80 Hz, gentle attenuation around 2-4 kHz) to reduce potential auditory fatigue. (3) Utilizing MIDI editing tools to modify or remove dissonant intervals (e.g., replacing tritones with perfect fourths or major thirds). (4) For timbre adjustment, specific virtual instrument plugins (e.g., Native Instruments' "The Gentleman" for piano, "Spitfire Audio" libraries for strings) were used to replace or enhance instrument sounds based on cancer type preferences. All adjustments were documented within the DAW project files for traceability.

(1) Tempo control: Rhythm speed was normalized into 60-88 BPM between two beats. Such tempo range was empirically associated with sedation and autonomic relaxation.

(2) Harmony avoidance: Dissonant intervals were strictly prohibited to avoid possible auditory stress or emotional agitation (especially tritone/diminished fifth, the most unstable interval).

(3) Timbre adjustment by cancer type: The composition instrumentation was moderately adjusted according to the patient's oncological classification. e.g. Piano notes were preferred for breast cancer patients; while natural white noise background without any artificial beat was added for lung cancer patients to give target patients a feeling of natural breathing and comfortable environment.

After the above human-tuning process, each piece of audio was further applied for real-time individuals' feedback collection. The real person subject rated the corresponding listening-emotion satisfaction on 5-points Likert scale from 1 to 5. The final list only adopted those compositions which gained feedback score of ≥ 4 . The efficiency of the optimization process was quantified by the number of required iterations (denoted as NNN) to reach the target satisfaction threshold. The following formula was used:

$$N = \frac{\text{Total Revisions}}{\text{Total Cases}}$$

The target value for manual optimization was set at $N \leq 1.5N \leq 1.5N$ iterations per patient, indicating a high degree of convergence between AI-generated outputs and patient preferences under therapist supervision.

2.2.3. Intervention duration and frequency design

Intervention protocol was delivered during an 4-week period, three times per week, with an interval of ≥ 24 h between sessions. Duration of each intervention was 30 min and delivered within 24 ± 2 h post-chemotherapy since this time period has been reported as psychologically vulnerable and optimal, in terms of neural plasticity, and emotional vulnerability. Music was delivered via Bose QuietComfort 45 active noise-canceling headphones, set to a fixed output volume of 50–60 dB to guarantee safe listening levels and high auditory quality. The temporal scheduling of the intervention was based on a chronotherapeutic model, taking into account patients' circadian physiology. Intervention was delivered in the morning window of the day (09:00–11:00) when cortisol levels are at their lowest, a time period in which increased emotional receptivity and neuroendocrine balance has been reported. The intervention dosage was operationalized as follows:

$$\text{Dose} = \sum_{i=1}^{12} (\text{Duration}_i \times \text{Adherence Rate}_i)$$

An adherence rate of $\geq 90\%$ was defined as the threshold for effective intervention exposure, ensuring sufficient dosage fidelity for outcome attribution.

2.2.4. Standard music intervention (control group)

Participants in the control group received a standardized music intervention, which was designed to reflect common practices in music therapy for relaxation. This intervention utilized a fixed playlist of pre-composed, commercially available

relaxation music. The playlist consisted of 12 instrumental pieces selected from established therapeutic music libraries (e.g., selections from Marconi Union's "Weightless," and similar ambient, slow-tempo instrumental tracks). The music was characterized by a slow tempo (60–80 BPM), lack of strong rhythmic accents, consonant harmonies, and the absence of sudden dynamic changes or lyrics. The same delivery protocol was followed as in the experimental group (Bose QuietComfort 45 headphones, 50–60 dB, 30-min sessions, three times per week for four weeks, within 24 h post-chemotherapy). This standardized approach ensured consistency across control group participants while lacking the personalized elements of the experimental intervention.

2.3. Outcome evaluation metrics

2.3.1. Anxiety assessment scales

Because anxiety and depression were assessed with both specificity and sensitivity, we measured these psychological outcomes with two standardized self-report instruments: Self-Rating Anxiety Scale (SAS) and Hospital Anxiety and Depression Scale—Anxiety Subscale (HADS-A). Testing was performed at four time points: baseline (T0), week 2 (T1), week 4 (T2), and one month follow-up (T3). In this study, we defined SAS standardized score ≥ 50 as clinically significant anxiety and HADS-A score ≥ 8 as presence of anxiety with clinical significance. Internal consistency of both instruments was good, with Cronbach's alpha = 0.82 for Self-Rating Anxiety Scale and Cronbach's alpha = 0.78 for Hospital Anxiety and Depression Scale—Anxiety Subscale, respectively.

The percentage change in total score (Δ) for each scale, representing dynamic shifts in anxiety severity over time, was calculated using the following formula:

$$\Delta = \frac{\text{Pre-score} - \text{Post-score}}{\text{Pre-score}} \times 100\%$$

A therapeutic response was defined as a $\geq 30\%$ reduction in anxiety scale scores ($\Delta \geq 30\%$), whereas clinical remission required SAS < 50 and HADS-A < 8 , indicating resolution of anxiety to subclinical levels across both instruments.

2.3.2. Physiological indicators

Physiological assessment encompassed measurement of two validated biomarkers of stress and autonomic regulation: heart rate variability (HRV) and salivary cortisol. HRV was measured in the frequency domain as the low-frequency to high-frequency ratio (LF/HF), which provides information about the sympathetic-parasympathetic balance of the autonomic nervous system. Data collection occurred at a sampling frequency of 1000 Hz, and analysis was performed on 5-min resting state recordings obtained under controlled laboratory conditions. Salivary cortisol was collected during the early morning (06:00–08:00) when the HPA axis is at its circadian peak. Samples were assayed using a commercially available enzyme-linked immunosorbent assay (ELISA) protocol. The lower limit of assay sensitivity was 0.1 $\mu\text{g/dL}$, which is sufficient to capture even small endocrine variations. The HRV improvement rate (η) was calculated to quantify autonomic restoration over time, based on the following formula:

$$\eta = \frac{\text{LF/HF}_{\text{post}} - \text{LF/HF}_{\text{pre}}}{\text{LF/HF}_{\text{pre}}} \times 100\%$$

The target threshold for HRV improvement was defined as $\eta \leq -15\%$, indicating a clinically meaningful enhancement in parasympathetic activity, reflecting reduced autonomic stress load.

2.3.3. Patient-reported outcomes and satisfaction assessment

To assess subjective experience at the time of the intervention, we recorded a Visual Analog Scale (VAS, range 0–10 points) immediately after each session to assess the amount of anxiety relief gained during that listening session. We also collected narrative feedback in free text format. After the intervention, the narrative feedback was analyzed qualitatively by thematic content coding. We recorded adherence by integrating a digital music log system, which automatically recorded the length of each session. A session was considered valid if the listening time was longer than 25 min (corresponding to $\geq 83\%$ adherence). Overall satisfaction (S) was calculated by integrating VAS scores under a structured post-intervention questionnaire (Q -score; Q -score was the mean score on a 5-point Likert scale; Cronbach's $\alpha = 0.85$). The composite satisfaction score was defined as:

$$S = \frac{\sum_{i=1}^n \text{VAS}_i + \text{Qscore}}{n + 1}$$

Scores of $S \geq 7.5$ were classified as indicative of high satisfaction with the intervention.

2.4. Statistical analysis

All statistical analyses were performed using SPSS software (Version 26.0; IBM Corp., Armonk, NY, USA). Continuous variables were expressed as means $\pm$ standard deviations ($\bar{x} \pm s$). When variables were normally distributed, independent-samples t -test was used for the comparisons between groups; otherwise, the Mann–Whitney U test was applied. Within a group, i.e., pre- vs. post-intervention, when variables were normally distributed, paired-samples t -test was used; otherwise, the Wilcoxon signed-rank test was applied. Categorical data were presented as frequencies and percentages. Chi-square (χ^2) test or Fisher's exact test was applied for inter-group comparisons when cell counts were not sparse; otherwise, the Wilcoxon signed-rank test was applied. To explore the intervention effects at multiple time points over long-term, GEE model was applied to accommodate repeated measures and to estimate main time effect as well as group $\times$ time interaction effect. Pearson's correlation coefficient was applied for parametric data correlation analysis; Spearman's rank correlation was used for non-parametric data. Two-sided $P < 0.05$ was considered statistically significant. When multiple comparisons were involved, Bonferroni correction was applied to control the family-wise error rate. Effect size was calculated by Cohen's d for continuous outcome and Cramer's V for categorical variables. Multiple imputation method was applied for missing data, which greatly reduced the bias.

3. Results

3.1. Baseline characteristics and group comparability

A total of 152 cancer patients with clinically significant anxiety completed the full intervention protocol (AI-generated music plus manual optimization, $n = 76$; standardized music, $n = 76$). The demographic and clinical characteristics of the participants at baseline are shown in **Table 1**. The two groups showed strong comparability with respect to all characteristics. In terms of age distribution, the mean age of the intervention group was 52.3 ± 8.7 years and that of the control group was 53.1 ± 9.2 years; there was no significant difference between the two groups ($t = 0.532$, $P = 0.596$). With respect to the sex ratio, males accounted for 50.0% of the intervention group and 52.6% of the control group; there was no significant difference between the two groups ($\chi^2 = 0.211$, $P = 0.646$). There was no significant difference between the two groups in the distribution of cancer types: breast cancer (26.3% vs. 28.9%), lung cancer (25.0% vs. 23.7%), colorectal cancer (21.1% vs. 19.7%), and other types (27.6% vs. 27.6%) ($\chi^2 = 1.024$, $P = 0.795$). With respect to baseline psychological measures, there was no significant difference between the two groups in either SAS scores (58.4 ± 5.2 vs. 57.9 ± 4.8 ; $t = 0.621$, $P = 0.536$) or HADS-A scores (9.1 ± 1.8 vs. 8.9 ± 1.6 ; $t = 0.782$, $P = 0.436$). In addition, there was no significant difference between the two groups in terms of physiological indicators. Morning salivary cortisol levels were 3.2 ± 0.9 $\mu\text{g/dL}$ in the intervention group and 3.3 ± 0.8 $\mu\text{g/dL}$ in the control group; there was no significant difference between the two groups ($t = 0.714$, $P = 0.477$). The LF/HF ratio, which reflects the autonomic nervous system function, was 2.6 ± 0.7 vs. 2.5 ± 0.6 ; there was no significant difference between the two groups ($t = 0.923$, $P = 0.358$). Overall, the results demonstrated the effectiveness of the randomization procedure and ensured a solid baseline equivalence between groups, providing a valid basis for subsequent outcome comparisons.

Table 1. Baseline characteristics of the two groups ($x \pm s/n(\%)$).

Variable	AI + manual group ($n = 76$)	Control group ($n = 76$)	$t/\chi^2/Z$	<i>P</i> -value
Age (years)	52.3 ± 8.7	53.1 ± 9.2	0.532	0.596
Male/Female	38/38	40/36	0.211	0.646
Tumor type			1.024	0.795
- Breast	20 (26.3%)	22 (28.9%)		
- Lung	19 (25.0%)	18 (23.7%)		
- Colorectal	16 (21.1%)	15 (19.7%)		
- Others	21 (27.6%)	21 (27.6%)		
SAS Score	58.4 ± 5.2	57.9 ± 4.8	0.621	0.536
HADS-A score	9.1 ± 1.8	8.9 ± 1.6	0.782	0.436
Morning cortisol ($\mu\text{g/dL}$)	3.2 ± 0.9	3.3 ± 0.8	0.714	0.477
LF/HF ratio	2.6 ± 0.7	2.5 ± 0.6	0.923	0.358

3.2. Pre-post and between-group differences in anxiety scores

As shown in **Table 2**, both music intervention approaches, namely, standard music intervention and personalized music intervention generated by AI, significantly decreased anxiety in cancer patients. However, better clinical efficacy on all psychological outcomes was achieved by music generated by AI with manual optimization. Within-group analysis showed that intervention group's SAS score significantly decreased from its baseline mean of 58.4 to 42.1, representing a 27.9% reduction ($P < 0.001$). Control group's SAS score significantly decreased from its baseline mean of 57.9 to 49.6, representing a 14.3% reduction ($P < 0.001$). Similar trends were observed for HADS-A scores, with 37.4% and 20.2% reductions, respectively, both significantly ($P < 0.001$). Between-group comparisons also supported the above findings. The mean SAS score of intervention group was 7.5 points lower than that of control group after intervention, namely, 42.1 points and 49.6 points, respectively ($P < 0.001$). Similar trends were observed for HADS-A scores, namely, 1.4 points (5.7 points vs. 7.1 points), representing 37.4% and 20.2% reductions, respectively ($P < 0.001$). All of the above differences were statistically significant and clinically meaningful. From clinical threshold perspective, the mean SAS score of intervention group was below the cutoff of 50, indicating clinical anxiety status was converted to subclinical condition after intervention. The mean SAS score of control group was still above the cutoff of 50, indicating limited anxiety resolution after intervention. All above results indicated that additional personalized value brought by AI-generated music intervention was statistically and clinically significant when manually refined by experts. From clinical threshold perspective, intervention group's mean SAS score was below the cutoff of 50, indicating clinical anxiety status was converted to subclinical condition after intervention. The mean SAS score of control group was still above the cutoff of 50, indicating limited anxiety resolution after intervention. All above results indicated that additional personalized value brought by personalized music intervention generated by AI when manually refined by experts was statistically and clinically significant. Our results offered strong empirical evidences to support the implementation of intelligent music therapy in comprehensive oncological care.

Table 2. Comparison of anxiety scores before and after intervention ($\bar{x} \pm s$).

Group	SAS score (pre)	SAS score (post)	Δ SAS (%)	HADS-A (pre)	HADS-A (post)	Δ HADS-A (%)	t/Z (within)	P-value (within)	t/Z (between)	P-value (between)
AI + Manual	58.4 $\pm$ 5.2	42.1 $\pm$ 4.3*	-27.9	9.1 $\pm$ 1.8	5.7 $\pm$ 1.2*	-37.4	12.74	<0.001	4.92	<0.001
Control	57.9 $\pm$ 4.8	49.6 $\pm$ 4.1*	-14.3	8.9 $\pm$ 1.6	7.1 $\pm$ 1.4*	-20.2	8.65	<0.001	-	-

* $P < 0.001$ vs. Pre-intervention within group.

3.3. Improvements in physiological indicators

As shown in **Table 3**, the music group with AI synthesis and manual tuning had significantly larger decreases in both autonomic nervous system function and endocrine biomarkers compared with the control group. Regarding salivary cortisol, morning cortisol levels decreased to 2.1 $\pm$ 0.6 μ g/dL from 3.2 $\pm$ 0.9 μ g/dL, indicating a 34.4% reduction ($P < 0.05$). This decrease was much more pronounced than that in

the control group, in which cortisol levels decreased to 2.8 ± 0.7 $\mu\text{g/dL}$ from 3.3 ± 0.8 $\mu\text{g/dL}$, showing a 15.2% reduction ($P < 0.05$). Therefore, the modulation of the HPA axis activity was much more evident in the personalized intervention group than in the control group, possibly because of the modulation of the hypothalamic-pituitary-adrenal axis. HRV analysis also showed markedly differentiated results between the two groups. The LF/HF ratio was reduced to 1.8 ± 0.5 from 2.6 ± 0.7 ($P < 0.05$), reflecting attenuation of the sympathetic nervous system. At the same time, the Root Mean Square of Successive Differences (RMSSD), which is an index of parasympathetic activity, was enhanced to 38.7 ± 7.5 ms from 28.4 ± 6.2 ms ($P < 0.05$). The RMSSD and LF/HF ratio also showed slight improvements in the control group ($P < 0.05$); however, the magnitude of change was much smaller than that in the intervention group. The objective biometric results obtained in this study are highly consistent with the psychological results reported earlier. The convergence of biometric and self-report results further confirms the effectiveness of personalized music with AI technology in regulating the neuroendocrine and autonomic function of cancer patients.

Table 3. Comparison of physiological indicators before and after intervention ($x \pm s$).

Group	Morning cortisol ($\mu\text{g/dL}$)	LF/HF ratio	HRV RMSSD (ms)			
	Pre	Post	Pre	Post	Pre	Post
AI + manual	3.2 ± 0.9	$2.1 \pm 0.6^*$	2.6 ± 0.7	$1.8 \pm 0.5^*$	28.4 ± 6.2	$38.7 \pm 7.5^*$
Control	3.3 ± 0.8	$2.8 \pm 0.7^*$	2.5 ± 0.6	$2.2 \pm 0.6^*$	27.9 ± 5.8	$32.1 \pm 6.3^*$

* $P < 0.05$ vs Pre-intervention within group.

3.4. Differential response by tumor stage

Mean percentage reduction of SAS scores in patients with early- (I–II) and advanced-stage (III–IV) malignancies showed meaningful differences in intervention response as demonstrated in **Table 4**. Participants with early-stage malignancies (Stage I–II) exhibited significantly larger mean percentage reductions of anxiety than participants with malignancies of advanced stage (Stage III–IV) in both intervention conditions. In the music AI composition plus manual fine-tuning group, mean percentage reduction of SAS scores was $32.5\% \pm 6.8\%$ in patients with Stage I–II malignancies and $24.1\% \pm 5.3\%$ in patients with Stage III–IV malignancies ($P < 0.001$). A similar trend was also found in the standardized music group, in which SAS scores decreased by $18.7\% \pm 5.2\%$ in patients with Stage I–II malignancies and by $11.9\% \pm 4.6\%$ in patients with Stage III–IV malignancies ($P < 0.001$). Regardless of cancer stage, intervention group participants exhibited significantly greater mean percentage reduction of anxiety compared with control group participants. The intergroup difference was significant among patients with Stage I–II malignancies ($t = 4.21$, $P < 0.001$) and among patients with Stage III–IV malignancies ($t = 3.87$, $P < 0.001$). Overall, these results suggest that although disease severity may represent a factor that moderates therapy response, individuals at different stages of cancer are clinically benefited by personalized music intervention with AI and expert fine-tuning across the disease continuum. Importantly, patients in earlier stages of cancer may provide greater psychological benefit, possibly because of higher level of

physiological resilience, less symptom burden, or greater attention to the therapeutic stimulus.

Table 4. Anxiety improvement in patients with different tumor stages (Δ SAS%, $x \pm s$).

Tumor stage	AI + Manual Group ($n = 76$)	Control group ($n = 76$)	t/Z	P -value
Stage I-II	-32.5 ± 6.8	-18.7 ± 5.2	4.21	<0.001
Stage III-IV	-24.1 ± 5.3	-11.9 ± 4.6	3.87	<0.001

3.5. Patient satisfaction and adherence outcomes

Patient-reported outcome data are presented in **Table 5**. Results show that the satisfaction and adherence level of intervention group patients in terms of music selected by AI and optimized by manual effort were significantly higher than those of intervention group patients who received standardized music. The mean satisfaction score of intervention group patients was 8.7 ± 1.2 , which was significantly higher than 6.3 ± 1.5 in the control group ($P < 0.05$). In terms of treatment adherence, $92.4 \pm 5.8\%$ of intervention group patients successfully completed all 12 scheduled sessions, whereas $83.1\% \pm 7.2\%$ of control group patients did so ($P < 0.05$), indicating higher treatment engagement. Regarding the preferences for musical elements, 68% of intervention group patients preferred compositions based on piano playing, and 72% preferred the mixing of natural sounds with music. In contrast, control group patients mostly favored classical music (54%) and light melodic structures (61%). Thematic analysis of the free-text feedback section also demonstrated the qualitative advantage of personalized interventions. The effectiveness of the AI's initial emotional matching was quantitatively assessed using cosine similarity. The average cosine similarity between the target emotional vector A (patient's state) and the feature vector B of the AI-generated music (before manual optimization) was 0.78 ± 0.09 (range: 0.61–0.92), indicating a generally high level of initial affective congruence. This result supports the AI model's capability in generating music that aligns well with the intended emotional profile, providing a solid foundation for the subsequent manual refinement process. Among intervention group patients, 82% explicitly stated that the music had “matched their emotional needs,” which was significantly higher than 43% of control group patients ($\chi^2 = 18.6$, $P < 0.001$). In summary, the results demonstrated the qualitative advantage of the AI-enhanced personalized music intervention model in terms of patient acceptance and treatment continuity. The high level of perceived personalization may play a key role in therapeutic engagement, which may explain the greater clinical efficacy demonstrated by the intervention across psychological and physiological outcomes.

Table 5. Patient satisfaction and adherence rates (%), $x \pm s$).

Group	Satisfaction (VAS 0-10)	Adherence rate (%)	Preferred music elements
AI + Manual	$8.7 \pm 1.2^*$	$92.4 \pm 5.8^*$	Piano (68%), nature sounds (72%)
Control	6.3 ± 1.5	83.1 ± 7.2	Classical (54%), light melody (61%)

* $P < 0.05$ vs. control group.

4. Discussion

4.1. Synergistic mechanisms of AI-generated music combined with human optimization

The combination of AI-generated music and manual fine-tuning by therapists manually resulted in a synergistic effect on reducing anxiety in cancer patients. This synergy can be decomposed into three closely related mechanisms: accuracy of emotional recognition, effectiveness of individual matching, and physiological effects of music. An affective generative model can encode input information, including emotional tags, prosody of voice, and stated musical preferences into high-dimensional latent spaces and generate music that is close to target emotional vectors. Such encoding and generation mechanisms can be learned based on large training corpora and deep learning representations [14]. Nevertheless, the generative model may still be generating musics with relatively uniform styles, limited emotional fluidity, and weak consideration to clinical context. These findings support the need for humans in the loop. The manual optimization process performed by board-certified music therapists were informed by clinical judgment, cultural considerations, and clinical contexts, and adjusted musical parameters, including tempo, timbre, and harmonic progressions. The human-in-the-loop process ensured that the generated musics are not only emotionally appropriate, but also physiologically contextually fitting to the circadian context, cognitive threshold, and patients' receptivity [15]. Importantly, we found that humans are playing a pivotal role in suppressing dissonant components and enhancing gradual emotional progression to compensate for the lack of contextual and humanistic consideration in the current generative model. Both the synergistic interaction between the algorithmic layer and therapeutic layer, and the interaction between precise individual matching and soothing stimulation, enhanced the immediate soothing stimulation of the intervention. More importantly, it enhanced the subjectivity of recognition and ownership of emotion, empowering patients to be more engaged in their own emotional regulation. We hypothesize that this process may stimulate parasympathetic activation and reduce cortisol secretion, which were observed in physiological outcomes. The combined effect of personalization and precision constitutes a new technology-human hybrid model of music intervention. This model demonstrated superior performance compared to standardized interventions commonly used in integrative psycho-oncological care. This model is scalable and easily replicable.

4.2. Neuroendocrine modulation: Physiological basis of music intervention effects

The therapeutic effectiveness of music therapy in oncology goes beyond providing comfort through emotional enrichment: it also derives from its dual regulatory actions on the autonomic nervous system (ANS) and the hypothalamic-pituitary-adrenal (HPA) axis. Anxiety states are often marked by increased sympathetic activity and enhanced cortisol release, both of which increase inflammatory burden, induce immunosuppression, and undermine treatment

compliance—potentially hastening disease development. To our knowledge, this is the first report showing that the AI-personalized music intervention reduced psychological stress through affective alignment and auditory resonance mechanisms, which subsequently triggered downstream neurophysiological changes [16,17]. That is, music with steady rhythm, soft timbre, and clear emotional valence is expected to activate the boundary regions of the amygdala and hippocampus while dampening activity of prefrontal cortical regions regulating higher emotions. Such neural engagement may underlie the inhibition of HPA axis hyperactivity, as assessed by morning cortisol in the present study. Additionally, the nucleus ambiguus and dorsal vagal complex are activated by auditory stimulation, which in turn improves parasympathetic tone (as assessed by increased RMSSD and decreased LF/HF ratio in heart rate variability) by reducing the overall physiological load from a stress-responsive state to a restorative one, consistent with the significant decreases in anxiety scores. The physiological efficacy of music as a non-pharmacological intervention resides in the capacity to exploit both the neuroplasticity of autonomic regulation and the programmability of emotional conditioning. In the case of personalized music, the fine-tuning of acoustic parameters (i.e., frequency, tempo, and texture) enables the nervous system to establish conditioned associative input and thus strengthen stable neural pathways for relaxation. The associated response output accumulates with repeated listening, conferring a sustained regulatory effect on the neuroendocrine system. This cumulative process improves the longitudinal stability of the intervention and augments its systemic coherence, endowing personalized music therapy with a biologically plausible and clinically scalable profile for the management of anxiety in cancer care.

4.3. Clinical implications of stage-dependent response to music intervention

The variability in therapeutic response across different stages of cancer highlights the interactive effects of disease burden and psychological receptivity to intervention. Patients diagnosed with early-stage malignancies (Stage I–II) were more responsive to the music intervention. Early-stage cancer is associated with shorter duration of disease, less damage to the body's resources, and a more plastic neuroendocrine system. Patients in this stage demonstrated higher affective responsiveness, greater ease of activation of the parasympathetic system, and thus were more capable of achieving larger and faster subjective changes in emotion through music. In addition, early-stage patients are generally in active treatment with good prognostic factors and a more positive psychological state and motivation to be more involved in self-regulatory activities. The state of physical preparedness combined with an open psychological state may create a positive feedback loop where the perceived benefits of an intervention amplify its actual effects, enhancing both subjective and physiological outcomes. In contrast, patients with later stages of cancer (Stage III–IV) have experienced prolonged physical suffering and treatment burden, often compounded by existential anxiety, leading to blunted affective responsiveness and dysregulation of the autonomic nervous system. These patients

may have higher thresholds for emotional engagement and in some cases, a state of resistance or rejection to engagement with stimuli. The therapeutic effects are still statistically significant but weaker [18]. The findings of the present study suggest that later-stage patients are still capable of experiencing clinically effective reductions in anxiety when exposed to individually tailored music interventions. In other words, the biological conditions that lead to efficacy are not necessary for reception but are closely related to emotional accessibility and cognitive readiness of the intervention content. These findings have important clinical implications. Stage-adapted intervention strategies should be explored, especially in later stages of illness, with an emphasis on increasing the readiness and openness of advanced patients to therapeutic engagement and emotional support.

4.4. The amplifying role of patient adherence and satisfaction

Adherence and satisfaction were strong mediators and magnifiers of treatment efficacy in a music-based intervention. In essence, these two variables measure the intensity of the intervention engagement and the level of psychological congruence with the intervention process. High adherence guarantees repeated exposure to the therapeutic stimulus (both in terms of frequency and duration), which is necessary for music to produce repeated regulatory effects on neural circuits and emotional states. Such repeated exposure is an absolute requirement for successful conditioning of relaxation and promoting systemic psychophysiological recovery. Satisfaction, in contrast, is the subjective validation of the intervention. It includes dimensions related to musical aesthetics, emotional resonance, and preference. When patients feel that the intervention is emotionally resonant with their state and meaningful to their self-concept, they are more likely to establish positive associations, maintain motivation, and produce internal emotional momentum that further reinforces the intervention outcome [19]. In our study, the extremely personalized music protocol (obtained after AI generation and fine-tuning by humans) led to increased satisfaction scores and adherence. This design addressed patients' aesthetic expectations and emotional needs, as well as the sense of control and therapeutic agency—two important psychological variables associated with health outcome. It is important to note that the success of any clinical intervention is not driven by the sophistication of the technology, but by the patient's level of engagement and emotional buy-in. Therefore, future applications of music therapy (especially if in digital or AI-enhanced form) should focus on designs that improve adherence and satisfaction [20]. This includes mechanisms that enhance control, preference, and engagement at the psychological and emotional level—including dimensions related to humanistic interaction, feedback, and personalized content delivery—which are crucial for the sustainability of the intervention model and treatment outcome.

4.5. Limitations and future directions

Although the encouraging results supported the feasibility and potential effectiveness of AI-generated music integrated with human fine-tuning to reduce anxiety in cancer patients, several limitations should be acknowledged, and further studies are expected to explore these issues. First, participants involved in the study

were from a single-center oncology hospital, which indicated that there was potential geographic and population-selection bias and limited external validity and generalization of the results. Although stratified randomization was used to ensure comparable baseline characteristics, large inter-individual heterogeneity in the emotional response to personalized music was still observed. The present study did not control for possible moderating variables, including personality, cultural background, and prior musical experience, which might have great impacts on the responsiveness of treatment [21]. Second, the intervention time was only four weeks. Although significant improvements in therapeutic effects were statistically achieved with this relatively short intervention, the lack of longitudinal follow-up after the intervention to observe the lasting effect of the therapy and possible relapse of symptoms was limited to our study, which hindered our understanding of the long-term effects of the intervention. Third, only HRV and salivary cortisol were selected as physiological markers, which reflected the characteristics of autonomic and endocrine responses and could not reflect the overall regulatory effects of neuroimmunology. To further understand the intervention's effects on the whole body, a more comprehensive physiological profiling method may be necessary [22,23]. Furthermore, while we defined the control intervention clearly in this revision, we acknowledge that the experimental intervention's novelty and the higher level of personalization and attention received by patients in the AI-hybrid group could have contributed to the outcomes through placebo or expectancy effects. The higher satisfaction and adherence in the experimental group might partly reflect this heightened engagement. Future studies could employ an active control group that receives a non-musical, similarly engaging and personalized intervention (e.g., personalized storytelling) to better isolate the specific effect of the personalized music itself. Further studies should focus on these limitations by using a larger sample from multiple centers to enhance representativeness and statistical power. Using integrated assessment frameworks including psychometric batteries, neuroimaging, and biomolecular markers may provide us with more insights into the mechanisms underlying the responses to personalized music and their neurophysiological substrates. The design of interventions may benefit from adaptive feedback systems that can dynamically adjust the contents of music according to the emotional states in a real-time manner, which may enhance the interactivity and ecological validity. Further, extended follow-up after the intervention is essential to observe the lasting effect of anxiolytic responses and its possible clinical outcomes over time. Finally, future studies should further explore the regulatory potential of music therapy on other comorbid psychological symptoms, such as depression and insomnia, and may apply it to other possible integrative modalities in the continuum of psycho-oncological care.

5. Conclusion

This study showed that a personalized music intervention protocol developed with the assistance of AI-generated music and fine-tuned by trained therapists, when compared to the specific standardized music intervention control used in this trial, can provide potential clinical advantages in the management of anxiety in cancer

patients. Compared with standard music interventions, the personalized intervention model yielded better intervention outcomes in anxiety-alleviation, physiological regulation (cortisol and HRV), and patients' satisfaction and adherence. The observed superiority suggests that the combination of AI-generation and human optimization is a promising approach, although the specific contribution of personalization versus the technological engagement itself warrants further investigation. By accurately controlling musical affective characteristics with respect to patients' conditions and applying professional therapeutic refinement, the present intervention model successfully realized the integration of technological accuracy and human wisdom, thereby enhancing emotional resonance, increasing psychological tolerance, and inducing parasympathetic activation, which stabilized the neuroendocrine profile. Subgroup analysis also indicated that early-stage cancer patients benefitted more, although clinically meaningful intervention effects were also found in late-stage cancer patients, highlighting the importance of stage-specific intervention strategies. The present findings suggested that personalized music therapy was safe, non-invasive, patient-centered, and promising as an adjunct in oncology, with high potential for clinical translation and application. Future studies should focus on increasing sample diversity, extending follow-up time, and further exploring how to realize optimal personalization according to feedback, in order to improve the intervention model and apply it to other emotional disorders and rehabilitation programs in oncology.

Author Contributions: HJ was responsible for all aspects of this work, including conceptualization, study design, methodology, data collection and analysis, writing—original draft, writing—review and editing, and final approval of the manuscript. The implementation of the study, resource acquisition, project administration, and funding acquisition (if applicable) were also carried out independently by the author. The author has read and agreed to the published version of the manuscript.

Funding: None.

Ethical approval: This study was reviewed and approved by the Institutional Review Board (Approval No.: 2022-085, as stated in the “2.1.1 Study Population” section of the main text). All procedures performed were in accordance with relevant ethical guidelines.

Informed consent statement: Written informed consent was obtained from all subjects involved in the study.

Acknowledgments: The author thanks Yonsei University in Seoul, South Korea, for providing academic support. Sincere gratitude is also extended to the patients and their families who participated in this study for their cooperation and trust, as well as to the medical staff of the oncology department at the cooperating hospital for their assistance during the implementation and coordination of the study.

Conflict of interest: The author declares no conflict of interest.

References

1. Luo T, Zhang S, Zhao M, et al. Efficacy of Music Therapy on Quality of Life in Cancer Patients: A Systematic Review and Meta-Analysis. *Psycho-Oncology*. 2025; 34(5). doi: 10.1002/pon.70165
2. Salgado-Vasco A, Torres-Morales J, Durán-Rojas CI, et al. The impact of group music therapy on anxiety, stress, and wellbeing levels, and chemotherapy-induced side effects for oncology patients and their caregivers during chemotherapy: a retrospective cohort study. *BMC Complementary Medicine and Therapies*. 2025; 25(1). doi: 10.1186/s12906-025-04837-7
3. Chen HH, Lai CH, Hou YJ, et al. The Efficacy of Music Intervention in Patients with Cancer Receiving Radiation Therapy: A Systematic Review and Meta-Analysis. *Cancers*. 2025; 17(4): 691. doi: 10.3390/cancers17040691
4. Li Y, Ji S, Tao Y, et al. The effect of music therapy on anxiety, depression, pain and sleep quality of lung cancer patients: a systematic review and meta-analysis. *Supportive Care in Cancer*. 2025; 33(3). doi: 10.1007/s00520-025-09213-2
5. Disha DJ, Bhavna CP, Sumathy S, et al. Harmonising Hope: Impact of Music Therapy on Cancer Pain and Palliative Care. *Indian journal of palliative care*. 2025; 31(1): 21-26.
6. Reyes Aragón L, Díaz AM, Suárez R, et al. Effect of a Single Live Group Music Therapy Intervention on Anxiety-State and Well-Being Levels During Chemotherapy: A Multicenter Randomized Clinical Trial Protocol. *Global Advances in Integrative Medicine and Health*. 2024; 13. doi: 10.1177/27536130241310241
7. Gezgin Yazıcı H, Ökten Ç, Karabulut E, et al. Effect of Music on Anxiety and Fatigue in Cancer Patients Undergoing Chemotherapy: A Randomized Controlled Trial. *Archives of Iranian Medicine*. 2024; 27(11): 611-617. doi: 10.34172/aim.31258
8. Agostinelli A, Denk T I, Borsos Z, et al. Musiclm: Generating music from text. *Arxiv preprint Arxiv*. 2023; doi: 10.48550/arXiv.2301.11325
9. Wei L, Yu Y, Qin Y, et al. From Tools to Creators: A Review on the Development and Application of Artificial Intelligence Music Generation. *Information*. 2025; 16(8): 656. doi: 10.3390/info16080656
10. Trigueros-Murillo A, Martinez-Calderon J, Casuso-Holgado MJ, et al. Effects of music-based interventions on cancer-related pain, fatigue, and distress: an overview of systematic reviews. *Supportive Care in Cancer*. 2023; 31(8): 488. doi: 10.1007/s00520-023-07938-6
11. Bradt J, Leader A, Worster B, et al. Music Therapy for Pain Management for People With Advanced Cancer: A Randomized Controlled Trial. *Psycho-Oncology*. 2024; 33(10). doi: 10.1002/pon.70005
12. Xu Z, Liu C, Fan W, et al. Effect of music therapy on anxiety and depression in breast cancer patients: systematic review and meta-analysis. *Scientific Reports*. 2024; 14(1). doi: 10.1038/s41598-024-66836-x
13. Yang HF, Chang WW, Chou YH, et al. Impact of background music listening on anxiety in cancer patients undergoing initial radiation therapy: a randomized clinical trial. *Radiation Oncology*. 2024; 19(1). doi: 10.1186/s13014-024-02460-3
14. Hyun JC, Ho S, Eu SC, et al. Music from noise-canceling headphones is beneficial against anxiety in male bladder cancer patients undergoing follow-up cystoscopy: A prospective randomized trial. *Urologia internationalis*; 2024.
15. Sweta M, Amit S, Jaya C, et al. Impact of Nada Yoga Music Therapy on Anxiety and Quality of Life in Ovarian Cancer Patients: A Randomized Controlled Trial. *International journal of yoga*. 2024; 17(2): 138-145.
16. Kobayashi M, Kako J, Kajiwara K. Response to “The effect of two different types of music played to cancer patients during chemotherapy on anxiety, nausea, and satisfaction levels.” *Supportive Care in Cancer*. 2024; 32(4). doi: 10.1007/s00520-024-08454-x
17. Dong J, Qu Y. Therapeutic Effect of Music Therapy on Patients with End-stage Cancer: A Retrospective Study. *Noise & health*. 2024; 26(121): 82-87.
18. Du Y, Guan Y, Thome KT, et al. Music Therapy and Music Intervention for NSCLC Patients Undergoing PET with Fear of Cancer Recurrence. *Integrative Cancer Therapies*. 2024; 23. doi: 10.1177/15347354241269898
19. Sun C, Sang S, Tang Y, et al. Effects of music therapy on anxiety in patients with cancer: study protocol of a randomised controlled trial. *BMJ Open*. 2023; 13(5): e067360. doi: 10.1136/bmjopen-2022-067360
20. Liou KT, McConnell KM, Currier MB, et al. Telehealth-Based Music Therapy Versus Cognitive Behavioral Therapy for Anxiety in Cancer Survivors: Rationale and Protocol for a Comparative Effectiveness Trial. *JMIR Research Protocols*. 2023; 12: e46281. doi: 10.2196/46281

21. Schandert L, Affronti ML, Prince M, et al. Music Intervention: Nonpharmacologic Method to Reduce Pain and Anxiety in Adult Patients Undergoing Bone Marrow Procedures. *Clinical Journal of Oncology Nursing*. 2021; 25(3): 314-320. doi: 10.1188/21.cjon.314-320
22. da Silva Santa IN, Schweitzer MC, dos Santos MLBM, et al. Music interventions in pediatric oncology: Systematic review and meta-analysis. *Complementary Therapies in Medicine*. 2021; 59: 102725. doi: 10.1016/j.ctim.2021.102725
23. Facchini M, Ruini C. The role of music therapy in the treatment of children with cancer: A systematic review of literature. *Complementary Therapies in Clinical Practice*. 2021; 42: 101289. doi: 10.1016/j.ctcp.2020.101289