Study protocol

Participants

All participants were recruited through a tinnitus community society website ("https://cafe.naver.com/onquest") and screened using the inclusion and exclusion criteria based on previous studies on sound therapy (Table 1) (Tyler et al., 2012; Tyler et al., 2020; Jin et al., 2021a). The participant selection process was conducted in a laboratory through an interview, and tinnitus was evaluated by a certified audiologist. Detailed information regarding the interview and the tinnitus evaluation is provided below in the "Outcome Measures" and "Study Design" sections. This study was approved by the institutional review board of Hallym University (HIRB-2020-069), and each participant received a written explanation regarding the aims, protocol, and procedures of the study. Each participant provided written informed consent before participation in the study.

The study participants were recruited from November 1, 2020 to February 28, 2021, and the follow-up examinations were performed from March 1, 2021 to June 1, 2021. The Consolidated Standards of Reporting Trials flow diagram for the clinical trial on sound therapy is shown in Figure 1. Ninety-two applicants were screened; 58 met the inclusion and exclusion criteria. The other 34 applicants were rejected for the following reasons: their hearing threshold levels at 0.5, 1, and 2 kHz were higher than the 40 dB hearing level (HL) (11 applicants); their average K-TPFQ scores were ≤30 points (10 applicants); they were currently receiving tinnitus-related rehabilitation or treatment (seven applicants); they had audiological problems (three applicants); or they were hearing aid users (three applicants). Among the 58 participants enrolled, 15 discontinued their participation during the study for the following reasons: 13 withdrew because of busy schedules and two felt that participation in the study was unhelpful. However, the analysis of the present study included discontinued participants' data in the intention-to-treat analysis (Fisher et al., 1990). Therefore, we analyzed the results of 58 participants (1-hour group: 20. 3-hour-group: 19, 5-hour group: 19) who were randomly allocated to one of the three groups, regardless of discontinued participation.

Stimulus and the Sound Therapy App

BBN, which is white noise ranging from 100 to 22,050 Hz, was used as the sound therapy stimulus; it was generated using an audio editing program (Adobe Audition 2018. Adobe Systems Complex, San Jose, CA, USA). The difference between the amplitudes of the BBN across frequencies was <1 dB. As the bandwidth and spectrum characteristics of the generated BBN can vary depending on the device used, this study was conducted using a single type of headphones (Etereo One, MIJ Co., Chuncheon, South Korea). The frequency range of these headphones was 20–20,000 Hz, which covers the average frequency range of adults.

To conduct sound therapy, a self-developed sound therapy application (app) was used (Jin et al., 2021b). This app allowed participants to perform sound therapy every day for a fixed amount of time. For example, if the app is set to 1 hour of use per day, the user can only perform sound therapy for up to 1 hour every day during the study period. In addition, the app was designed to specify the duration of a sound therapy session. For example, if the total daily sound therapy duration is 3 hours and the duration of one session is set to 1 hour, the participant should perform it three times on that day. This app was designed to maintain the type and

volume of the set sound source, and the participants were instructed to only use the function to turn the sound on and off through this app. Only the researchers could modify the settings of the app through a control program. The daily target time and duration of sound therapy performed were displayed on the app display. Therefore, the participants could easily check the time actually spent in sound therapy during the day. This app was designed in such a way that once the participants completed their daily sound therapy sessions, they could no longer perform sound therapy on that same day. The app reset again at 12:00 AM. Each participant's daily hours of sound therapy were recorded using the data logging system of the app. The researchers received each participant's usage duration of sound therapy in real-time through this app and measured each participant's actual duration of sound therapy using these data. In addition, the app's volume can increase from quiet to maximum intensity in 0–100 increments, while the volume of the general mobile phone has 15–20 increments only; therefore, the researchers could reach a level close to the participant's mixing point through a finer volume control than that allowed by a normal mobile phone.

Outcome Measures

In each group, the effectiveness of sound therapy was determined based on changes in the tinnitus loudness level, VAS for loudness score, and K-TPFQ score between the baseline and follow-up (3-month) time points. In addition, the Pearson's correlation coefficient was calculated to assess a linear association between the total hours of sound therapy actually performed and the degree of improvement of tinnitus for participants. The tinnitus loudness level was measured as dB SL in the participants' tinnitus frequency band, which is a well-established method for loudness matching tests (Henry, 2016). For instance, 10 dB SL meant that the tinnitus loudness level of the participant was 10 dB louder than their hearing threshold at that frequency band. The tinnitus loudness matching test was conducted along with a pitch-matching test. Pitch matching was performed by adjusting between 500 and 12,000 Hz at 500-Hz intervals. The participants listened to two comparative tones and chose the one they considered similar to their tinnitus sounds. After the pitch-matching test, loudness matching was performed to identify the point that the participants considered as similar to their tinnitus loudness, with the intensity being increased from 0 dB HL. The stimulus level increased at 2-dB intervals until the stimulus was just noticeable; subsequently, it increased at 1-dB intervals until loudness matching was achieved. Pitch and loudness matching were performed thrice; identical values measured two or more times were adopted as the result.

The VAS is a measurement tool that rates a continuum of values and seeks to measure a characteristic or attitude that is not easily measurable directly (Gould et al., 2001). In tinnitus studies, the VAS has been widely used to measure the subjective perception of tinnitus loudness or annoyance in people with tinnitus (Adamchic et al., 2012; Tass et al., 2012; Jin et al., 2021b). In the present study, the VAS score for loudness was measured and consisted of a 20-cm horizontal line with the left side corresponding to the absence of tinnitus (score of 0 points) with an image of a smiling face and the right side corresponding to the maximum imaginable loudness (score of 100 points) with an image of a frowning face (Jin et al., 2021b). The participants marked a spot on the line that indicated their present tinnitus loudness level.

The TPFQ is a tinnitus questionnaire that was developed to evaluate the severity of tinnitus by measuring the effect of tinnitus on sleep, hearing, emotion, and concentration (Tyler et al.,

2014). In the present study, the Korean version of the TPFQ (K-TPFQ) was used, whose validity and reliability have been confirmed (Shin et al., 2019; Heo & Jin, 2019). The K-TPFQ consists of five questions in each category (sleep, hearing, emotion, and concentration), for a total of 20 items. Each item is rated on a scale from 0 (strongly disagree) to 100 (strongly agree) points. For example, if a participant answers the item "I have difficulty getting to sleep at night because of my tinnitus" in the sleep category with a score of 100, it means that the participant completely agrees with that statement. Conversely, a score of 0 means that the participant completely disagrees with the statement. Therefore, higher the scores can be interpreted as representing greater difficulty due to tinnitus. The score for each category was calculated as the average score of all items in that category, and the total score was calculated as the average score of all items. In general, a decrease by ≥13 points of an individual's total K-TPFQ score or the K-TPFQ score for each category, which range from 0 to 100 points, can be interpreted as a significant effect; a total score of ≤30 points for an individual is interpreted as indicating that the adverse effect due to tinnitus is small (Tyler et al., 2014; Tyler et al., 2020; Jin et al., 2021a).

To assess a linear association between the total hours of sound therapy actually performed and the degree of improvement of tinnitus for participants, Pearson's correlation coefficient was calculated (Mukaka, 2012). Tinnitus improvement for each individual was calculated as the difference between the baseline and 3-month of tinnitus loudness level, VAS for loudness score, and K-TPFQ scores. In general, a correlation coefficient between -0.3 and 0.3 is interpreted as negligible correlation, between 0.3 and 0.5 (-0.3 and -0.5) as low positive (negative) correlation, between 0.5 and 0.7 (-0.5 and -0.7) as moderate positive (negative) correlation, between 0.7 and 0.9 (-0.7 and -0.9) as high positive (negative) correlation, and between 0.9 and 1.0 (-0.9 and -1.0) as very high positive (negative) correlation.

Study Design

The present study was conducted at the Research Institute of Audiology and Speech Pathology of Hallym University to confirm the effect of sound therapy according to the daily therapy hours. Therefore, counseling other than education related to the method of performing sound therapy during the study period was not provided, but counseling related to tinnitus was provided at the last visit, after the study was completed. The protocol for the present study is presented in Figure 2. A certified audiologist conducted the initial interview at the first visit to assess tinnitus duration, age, and inclusion and exclusion criteria. A puretone audiometry test and tinnitus evaluation (tinnitus loudness level evaluation, VAS for loudness, and K-TPFQ) were also performed at the first visit. The initial interview and measurements of VAS for loudness and K-TPFQ were conducted in a counseling room, and the puretone audiometry and tinnitus loudness level evaluations were conducted in a double-walled sound booth. The participants were screened based on the initial interviews and measurements of tinnitus and puretone audiometry. The initial interview was conducted in a 1:1 ratio with an audiologist for each participant. The audiologist asked the interview questions (i.e., tinnitus duration, age, and inclusion and exclusion criteria) and recorded the answers in the interview document. The audiologist also explained to the participants how to fill out the VAS for loudness and the K-TPFQ. In the counseling room (quiet condition), each participant read and filled in their response to the VAS for loudness and rated each item of the K-TPFQ. The audiologist waited in the room next to the counseling room and answered any questions the participants had. The

puretone audiometry and tinnitus loudness level tests were conducted using an audiometer (GSI Audiostar Pro, Grason-Stadler, Eden Prairie, MN, USA) and headphones (HDA 200, Sennheiser Electronic GmbH & Co, Wennebostel, Germany). The puretone audiometry test was performed to confirm the criteria for the inclusion of participants, the difference in the puretone average (PTA) of 500, 1,000, and 2,000 Hz between the groups at baseline, and the difference in hearing thresholds of each participant between baseline and the 3-month follow-up visit. A rest period of 5 min was provided between each test. In addition, an additional break time was provided whenever the participant wanted it. Based on the screening, the selected participants were randomly assigned to one of three groups according to the daily hours of sound therapy (1 hour [1-hr], 3 hours [3-hr], and 5 hours [5-hr]). Randomization using an Excel spreadsheet (Microsoft, Redmond, WA, USA) was performed for allocating participants to groups. This randomization program assigned numbers 1–58 to the 58 participants and randomly changed their order. Those assigned numbers 1–20 were allocated to the 1-hr group, those assigned numbers 21–39 were allocated to the 3-hr group, and the remaining participants were assigned to the 5-hr group. The total time for conducting the initial interview and evaluating hearing thresholds and tinnitus ranged from 1.5 to 2 hours for each participant.

At the end of the initial visit, the audiologist conducted a session on how to use the sound therapy app with headphones. The headphones (Etereo One, MIJ Co.) and the app were provided for free. The main contents of the session were information pertaining how to connect the headphones and the sound therapy app, turning the app on and off, and what the screen on the app indicated. The volume of the sound therapy stimulus (BBN) was adjusted close to the mixing point level of each participant and set based on the most responsive mixing point level measured using the ascent method five times in the counseling room. The participants were instructed as follows: “You will now start listening to an external sound stimulus through the headphones. At first, you will probably hear tinnitus, but not the sound stimulus. As I adjust the volume of sound stimulus, I will ask you whether you can hear the sound stimulus. You just have to tell me if you can hear the sound stimulus or not. The volume at which the external sound stimulus starts to be heard while tinnitus is still heard is the suitable volume of the external sound stimulus.” For example, if the volumes that a participant reported as the mixing point were 14, 13, 14, 15, and 14, then 14, which had the highest frequency, was determined as the sound therapy volume of that participant. The audiologist confirmed whether the final determined volume of the app was blended with the participant’s tinnitus by asking the participant. The participants were instructed to perform sound therapy when their tinnitus bothered them and to proceed with sound therapy for 1 hour at a time to keep the timing and pattern of sound therapy the same among the participants. Even if annoying tinnitus sounds were perceived multiple times a day, the audiologist instructed the participants to proceed with sound therapy only for a set number of times per day. In the 1-hr, 3-hr, and 5-hr groups, the participants performed sound therapy once (1 hour), thrice, and five times a day, respectively. As all the participants always perceived tinnitus sound in a quiet environment, none of the participants reported that sound therapy was performed when the participant did not perceive tinnitus. However, among the participants in the 3-hr and 5-hr groups, there were cases where sound therapy was performed continuously for 2 or 3 hours. The participants reported that tinnitus was perceived even after 1 hour of sound therapy. Therefore, they performed additional sound therapy. In addition, participants were guided to perform sound therapy in a space similar to the environment in which the volume of sound therapy was determined (counseling room)

and to check whether the sound therapy stimulus was blended with tinnitus sound (mixing point level) when they started the sound therapy. The participants were guided if they could not hear at the mixing point level, at which they conducted the sound therapy and were asked to adjust the sound therapy volume over the phone or during a visit. At the initial visit, participants were informed that they could discontinue the sound therapy at any time. The time taken for the education session was between 30 and 45 minutes.

After the initial visit, the participants performed daily sound therapy for 3 months. In some previous studies, a duration of 3 months was considered as the minimum duration to confirm the long-term effects of sound therapy (Henry et al., 2006; Jin et al., 2021b). Therefore, in the present study, a 3-month period was set as the follow-up duration. If the participants did not complete their daily sound therapy sessions, the researchers sent text messages encouraging them to proceed with sound therapy the next morning. After completing the 3-month sound therapy, participants visited the laboratory for measurements of puretone audiometry, tinnitus loudness level evaluation, VAS for loudness, and K-TPFQ. The puretone audiometry test was performed to measure changes in hearing during the 3 months of sound therapy, and none of the participants showed a change >5 dB. The tinnitus loudness level test was performed in the same manner as in the initial visit. The participants completed the VAS for loudness and K-TPFQ by themselves, alone in the counseling room, to avoid biases. The time taken for measurements ranged from 50 to 60 minutes. The results of the sound therapy of each participant were interpreted at the end of each session. Fifteen participants who discontinued the study visited the laboratory at 3 months to return the headphones and receive the participant rewards, and, like the other participants, performed measurements of puretone audiometry, tinnitus loudness level evaluation, VAS for loudness, and K-TPFQ. Pearson's correlation coefficients between actually performed total hours of sound therapy and degree of improvement of tinnitus were calculated by the researchers after completing the data collection.

Data Analysis

Baseline comparisons of the mean PTA, duration of tinnitus, age, tinnitus frequency and loudness level (dB SL), VAS score for loudness, and K-TPFQ score among the three groups (1-hr, 3-hr, and 5-hr) were performed using one-way analysis of variance (ANOVA). To identify interactions between the three groups (1-hr, 3-hr, and 5-hr) and the two time points (baseline and 3-month follow-up) in the tinnitus loudness level, VAS scores for loudness, and K-TPFQ scores, repeated measures ANOVA was used. When a significant interaction was observed, a pairwise comparison was performed to identify the group with a significant change of outcome measures according to time points. Repeated measures ANOVA uses the estimated average value, but the actual and estimated average values of this study were identical; therefore, the average values were reported when describing the results. All statistical analyses were performed using SPSS software (version 25.0; IBM Corp., Armonk, NY, USA), and an error threshold alpha=0.05 was used as the cut-off for significance.

Additionally, to assess the linear association between total hours of sound therapy and the degree of tinnitus improvement according to outcome measures, Pearson's correlation coefficient was calculated.