

---

**Supplementary information**

---

# **Multicentre gene therapy for *OTOF*-related deafness followed up to 2.5 years**

---

In the format provided by the  
authors and unedited

# **Multicenter gene therapy for OTOF-related deafness followed up to 2.5 years**

Luoying Jiang, Xiaoting Cheng, Jun Lv, Yuxin Chen, Xiaoyun Chen, Rongqun Zhai, Liqin Zhang, Lei Han, Yiling Zhang, Juhong Zhang, Di Deng, Zhicheng Huang, Qi Cao, Xin Zhang, Daqi Wang, Yizhe Wang, Liheng Chen, Sha Yu, Luo Guo, Bowen Zhang, Hui Wang, Yi Zhou, Liling Dai, Wei Wang, Longlong Zhang, Yanbo Yin, Guiqing Cheng, Ziyi Zhou, Wuqing Wang, Bing Chen, Wei Lu, Hongqun Jiang, Zhiqiang Gao, Dazhi Shi, Yuanping Xiong, Yu Zhao, Wei Yuan, Qin Wang, Guodong Feng, Huawei Li, Zheng-Yi Chen, Yilai Shu

## CONTENTS

|                                                                                                                                                                      |           |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| <b>Supplementary Notes.....</b>                                                                                                                                      | <b>3</b>  |
| The list of clinical trial centers.....                                                                                                                              | 3         |
| <b>Supplementary Methods .....</b>                                                                                                                                   | <b>4</b>  |
| Surgery procedure .....                                                                                                                                              | 4         |
| Auditory brainstem response .....                                                                                                                                    | 4         |
| Auditory steady-state response .....                                                                                                                                 | 5         |
| Behavioral audiometry.....                                                                                                                                           | 6         |
| Distortion product otoacoustic emission.....                                                                                                                         | 7         |
| Auditory and speech questionnaires and tests .....                                                                                                                   | 7         |
| Immune response and vector shedding .....                                                                                                                            | 8         |
| Follow-up for endpoints.....                                                                                                                                         | 9         |
| Supplementary statistical analysis .....                                                                                                                             | 9         |
| <b>Supplementary Tables.....</b>                                                                                                                                     | <b>11</b> |
| Supplementary Table 1   Baseline characteristics and follow-up duration of participants by center after expansion to multicenter period <sup>a</sup> .....           | 11        |
| Supplementary Table 2   Median (Range) for the cholesterol levels, lymphocyte counts, and other blood counts.....                                                    | 12        |
| Supplementary Table 3   Sensitivity analysis of dose-dependent effects on ABR thresholds in participants with hearing recovery .....                                 | 13        |
| Supplementary Table 4   Sensitivity analysis of dose-dependent effects on ASSR thresholds in participants with hearing recovery .....                                | 14        |
| Supplementary Table 5   Sensitivity analysis of dose-dependent effects on behavioral thresholds in participants with hearing recovery .....                          | 15        |
| Supplementary Table 6   Sensitivity analysis of the relationship between hearing thresholds and IT-MAIS/MAIS scores in participants with hearing recovery.....       | 16        |
| Supplementary Table 7   Sensitivity analysis of the relationship between hearing thresholds and SSQP-noise related scores in participants with hearing recovery..... | 17        |
| Supplementary Table 8   Sensitivity analysis of the relationship between hearing thresholds and MUSS scores in participants with hearing recovery.....               | 18        |
| <b>References.....</b>                                                                                                                                               | <b>19</b> |

## **Supplementary Notes**

### **The list of clinical trial centers**

**The clinical trial was a multi-center trial, conducted at 8 sites in China, including:**

Center 1: Eye & ENT Hospital of Fudan University (Shanghai, China)

Center 2: Peking Union Medical College Hospital (Beijing, China)

Center 3: The Second Xiangya Hospital of Central South University (Changsha, Hunan, China)

Center 4: Chongqing General Hospital (Chongqing, China)

Center 5: West China Hospital, Sichuan University (Chengdu, Sichuan, China)

Center 6: The First Affiliated Hospital of Nanchang University (Nanchang, Jiangxi, China)

Center 7: The First Affiliated Hospital of Zhengzhou University (Zhengzhou, Henan, China)

Center 8: The Second Hospital of University of South China (Hengyang, Hunan, China).

## **Supplementary Methods**

### **Surgery procedure**

The procedure was conducted under an endoscope (7220AA, Karl Storz). A 5-mm circumferential incision ( $200^{\circ}$ – $270^{\circ}$ ) around the tympanic membrane was made to elevate the tympanomeatal flap anteriorly and superiorly, followed by lifting the tympanic membrane with a crochet needle. The chorda tympani nerve was identified and dissected carefully. To expose the stapes footplate and round window, a 2-mm drill was used to remove part of the posterior bony wall of the external auditory canal, with additional resection of the lateral canal wall of the attic and antrum. The fustis, as a key landmark, guided the localization of the round window. A 1-mm drill (Medtronic) was employed to remove the anterior wall of the round window niche, exposing the round window membrane (RWM) injection site. The stapes footplate was fenestrated using a 0.5-mm drill (Medtronic) with a three-edged needle. AAV1-hOTOF was injected through a custom microneedle connected with a microliter injection system. Post-injection, the RWM and footplate were sealed with fascia, and the flap was repositioned.

### **Auditory brainstem response**

Auditory brainstem response (ABR) measurements were conducted under controlled neurophysiological conditions to determine hearing thresholds. Evaluations were performed in an acoustically shielded chamber with subjects under sedation. A standardized three-electrode system was implemented, with the active electrode positioned at the vertex region of the scalp (Cz), the reference electrode on the ipsilateral mastoid, and the ground electrode secured to the contralateral mastoid process. The high-pass and low-pass filters for ABR are set to 1500 Hz and 100 Hz, respectively. The stimulus rate is 29.9 per second, with a total of 1024 sweeps at minimum. The masking type is white noise at 55 dB nHL at minimum when necessary. Insert earphones are used as the transducer, and the stimulus polarity is alternating for ABR. Sounds were delivered via headphones, and stimulation protocols incorporated broadband clicks and frequency-specific tone burst stimuli at 0.25, 0.5, 1, 2, and 4 kHz using an auditory evoked potential system. Bio-logic Navigator PRO (CA, USA) was used in the Eye & ENT Hospital of Fudan University, Peking Union Medical College Hospital, West China Hospital, The First Affiliated Hospital of Nanchang University,

and The Second Hospital of University of South China; while Interacoustics Eclipse (Middelfart, Denmark) was used in The Second Xiangya Hospital of Central South University, Chongqing General Hospital, and The First Affiliated Hospital of Zhengzhou University. Waveform identification required repeatable wave V morphology, and the hearing threshold was determined as the lowest level identified by two audiologists.

### **Auditory steady-state response**

Auditory steady-state response (ASSR) testing was implemented through an auditory evoked potential system in a dual-chamber sound-isolated suite, following an established methodology. Bio-logic Navigator PRO (CA, USA) was used in the Eye & ENT Hospital of Fudan University, Peking Union Medical College Hospital, Chongqing General Hospital, West China Hospital, The First Affiliated Hospital of Nanchang University, and The First Affiliated Hospital of Zhengzhou University, and The Second Hospital of University of South China; while Interacoustics CHARTR EP200 (Middelfart, Denmark) was used in The Second Xiangya Hospital of Central South University. Frequency-specific thresholds were evaluated via air-conducted stimuli across five frequencies (0.25, 0.5, 1, 2, 4 kHz) with intensity modulation ranging from 20-120 dB nHL in 5-10 dB nHL increments. The modulation frequencies for the left ear corresponding to each carrier frequency are 80, 82, 87, 91, and 96 Hz, respectively; for the right ear, they are 81, 84, 89, 94, and 99 Hz, respectively. The acceptance criterion for the *p*-value is less than 0.049. Artifact rejection is performed when the noise level is  $\geq 40$  microvolts. A three-electrode configuration was employed: active electrode at Fz (mid-frontal position), reference at the ipsilateral mastoid, and ground at Fpz (nasion region). The amplifier was set to a gain of 100,000, featuring cutoff frequencies of 10 Hz and 300 Hz. The signals were digitized at a sampling interval of approximately 1.37 ms. Each measurement was an average of approximately 20–24 epochs. Data acquisition involved 3-minute recording epochs with trial averages per intensity level. Automated response detection algorithms validated ASSR thresholds at 99% confidence criteria, ensuring objective signal identification without manual waveform interpretation.

## **Behavioral audiometry**

Behavioral audiometry includes conventional audiometry, play audiometry (PA), visual reinforcement audiometry (VRA), and behavioral observation audiometry (BOA). The specific method is selected based on the participant's age, and the final choice depends on the participant's level of cooperation. In general, VRA is used for participants younger than 2 years, and BOA is performed if VRA cannot be completed<sup>1</sup>. PA is used for participants aged 2 to 5 years, and conventional audiometry can be attempted for those older than 5 years<sup>1</sup>.

Conventional audiometry was conducted in an acoustically isolated chamber using a clinical diagnostic system with Madsen Conera clinical audiometer (Otometrics, Taastrup, Denmark) and TDH-39 audiometric earphones (Telephonics, Tarmingdale, NY, USA). The stimulus intensity at each frequency (0.125-8 kHz) began at 100 dB HL. Cooperative subjects underwent air-conduction testing through a modified Hughson-Westlake protocol: Initial stimuli at 100 dB HL decreased in 10 dB HL decrements until non-response, followed by 5 dB HL increments. The tested frequencies were 0.125-8 kHz. The threshold was defined as the lowest sound level at which a consistent response could be elicited. The bone-conduction testing was conducted through a similar manner with a bone vibrator placing at the mid of the forehead.

For PA, the audiologist first builds trust and conditions the child to place a designated toy into a box upon hearing sound, with correct responses rewarded verbally or via gestures. In formal testing, sounds are presented irregularly with several-second intervals, and auditory sensitivity is assessed through the child's consistent play-based responses. These test was performed using the audiostar pro (Grason-Stadler, USA). For VRA, before testing, conditioned reflexes are established through repeated training—immediate visual rewards (e.g., light-up toys) are given when the infant turns toward the sound. During testing, the infant sits facing the loudspeaker at a 45-90° angle; visual reinforcers in a dark box become visible only with sound, and head-turn responses are recorded to estimate hearing thresholds. For BOA, warble tone is used as the stimulus; after sound presentation, both attentive-type, orienting behaviours and reflexive responses from the infants are carefully observed, such as eyelid reflex and head turning toward the sound source, to evaluate auditory responsivity of the child. In all behavioral testing, tones are given through earphones if the participant tolerates, or

through the sound-field speaker if not. Operant conditioning procedures are done before formal testing when necessary.

### **Distortion product otoacoustic emission**

Distortion product otoacoustic emission (DPOAE) was assessed in an acoustically controlled chamber using a AudX Plus OAE analyzer (Madsen). Pre-test procedures included otoscopic examination to ensure unobstructed ear canals. Binaural measurements employed dual primary tones ( $f2/f1$  ratio = 1.22), with the lower-frequency primary tone at 65 dB SPL and the higher-frequency primary tone at 55 dB SPL. Frequency sweeps targeted  $f2$  frequencies at 0.5, 1, 2, 4, and 8 kHz. The  $2f1-f2$  distortion product amplitude and ambient noise floor were systematically recorded across test frequencies. Data validity required the signal-to-noise ratio (SNR) to be  $\geq 6$  dB SPL, with automated analysis confirming response presence. The number of present DPOAEs at baseline (“Pre”) was defined as the number of frequencies (out of the five tested) that met the SNR  $\geq 6$  dB SPL criterion and were automatically determined to have a valid response.

### **Auditory and speech questionnaires and tests**

Speech perception was assessed through questionnaires—Infant-Toddler Meaningful Auditory Integration Scale/Meaningful Auditory Integration Scale (IT-MAIS/MAIS)<sup>2</sup>, Meaningful Use of Speech Scale (MUSS)<sup>3</sup>, The Speech of the Speech, Spatial, and Other Qualities of Hearing Scale for Parents (SSQ-P)<sup>4,5</sup>, Categories of Auditory Performance (CAP)<sup>6</sup>, Speech Intelligibility Rating (SIR)<sup>7</sup>, and speech tests were conducted using the Mandarin Speech Perception<sup>8,9</sup>.

The IT-MAIS/MAIS results presented in this study refer to the sum of the scores for questions 3 to 10, as questions 1-2 specifically address hearing assistive device usage and were excluded from the analysis. Hence, the remaining 8 items result in a total IT-MAIS/MAIS score range of 0-32. The MUSS comprises 10 items, each scored on a 0 – 4 scale (0 = never, 1 = rarely, 2 = occasionally, 3 = frequently, 4 = always) with a maximum score of 4 per item, resulting in a total possible score of 40.

The SSQ-P comprises three dimensions: Speech Perception (9 items), Spatial Hearing (6 items), and Auditory Quality (8 items). Each domain scores out of 10 points. For the purpose of this study, we extracted noise-related items focusing on speech

perception in noisy environments, including Items 1, 4, 5, 6, and 8 from the Speech Perception dimension and Item 2 from the Auditory Quality dimension.

The scores of the CAP range from 0 to 7, where each score corresponds to a specific auditory ability: 0 = No awareness of environmental sounds; 1 = Awareness of environmental sounds; 2 = Responds to speech sounds (e.g. "go"); 3 = Identification of environmental sounds; 4 = Discrimination of some speech sounds without lip-reading; 5 = Understanding of common phrases without lip-reading; 6 = Understanding of conversation without lip-reading; 7 = Uses of telephone with known listener.

The scores of the SIR range from 1 to 5, where each score corresponds to a specific speech intelligibility criterion: 1 = Connected speech is unintelligible. Pre-recognisable words in spoken language; 2 = Connected speech is unintelligible. Intelligible speech is developing in single words when context and lipreading clues are available; 3 = Connected speech is intelligible to a listener who concentrates and lipreads; 4 = Connected speech is intelligible to a listener who has little experience of a deaf person's speech; 5 = Connected speech is intelligible to all listeners.

Speech perception tests included monosyllable, disyllable, and sentence in quiet environments. The speech reception threshold of monosyllable and disyllable recognition was evaluated in noise environments.

### **Immune response and vector shedding**

To monitor the immune response, the anti-AAV1 neutralizing antibody was detected. For neutralizing antibody testing, blood samples were collected and preprocessed at the respective centers. Subsequently, processed samples were uniformly shipped to the specialized institution for standard laboratory tests, according to our previous reports<sup>10,11</sup>. HEK293FT cell line with no mycoplasma contamination was used to detect the anti-AAV1 neutralizing antibody. Authentication for HEK293FT cell line was validated by STR profiling (NANJING COBIOER BIOSCIENCES CO., LTD.). Circulating T cell responses to the AAV1 capsid were detected by interferon-gamma enzyme-linked immunosorbent spot (ELISpot) assay using an ELISpot PRO: Human IFN- $\gamma$  (ALP) Kit (MABTECH, code: 3420-2AST-10); to detect AAV vector shedding in blood, a quantitative polymerase chain reaction was performed, according to previous reports<sup>10,11</sup>.

### **Follow-up for endpoints**

Auditory testing, including ABR, ASSR, behavioral audiometry, and DPOAE, and speech perception of questionnaires and tests were assessed at baseline, 4, 6, 13, 26, and 52 weeks, as well as 1.5, 2, and 2.5 years. Blood and urine routine, blood biochemistry, coagulation function, and clinical manifestations were monitored at baseline, days 3 and 7, weeks 2, 4, 6, 13, 26, and 52, and years 1.5, 2, and 2.5. Hearing and speech perception data collected only at week 4 were combined with the week 6 data for analysis. Anti-AAV1 neutralizing antibodies were detected at baseline, day 3, weeks 2, 6, 13, 26, and 52, and years 1.5, 2, and 2.5. T cell responses to AAV capsid were evaluated by ELISpot assays at baseline and week 13. The level of vector DNA was assessed at baseline and within 1 week after treatment (day 3 or 7). The longest follow-up time point is scheduled 5 years after gene therapy.

### **Supplementary statistical analysis**

Baseline characteristics and safety analyses were performed on all enrolled participants, while efficacy analyses were performed on participants with hearing recovery. Two participants discontinued follow-up for personal reasons (dropout rate, 4.8%). One discontinued after 1.5 years, and the other after 26 weeks. The missing data were considered missing at random. At baseline, the thresholds of no response at the maximum sound intensity level were recorded as “>”. As only a few participants exhibited non-responses at certain frequencies during follow-up, the “>” was not used post-treatment. Similarly, post-treatment responses below the minimum measurable threshold were also not specifically marked. For analysis, both types of thresholds were calculated as the corresponding maximum or minimum sound intensity values.

Exploratory analyses of factors associated with auditory and speech outcomes were conducted using the linear mixed-effects model (LMM) with an exchangeable covariance structure in participants with hearing recovery. Sensitivity analyses and subgroup analyses were also performed. For the sensitivity analyses, adjusted LMM was used to identify influential factors, while simple models were used to assess robustness. For hearing thresholds, the simple model included the average hearing thresholds (ABR, ASSR, or behavioral audiometry), dose group, and follow-up time points, while the adjusted model additionally accounted for potential confounders, including age group, the number of present DPOAE at baseline, *OTOF* variants, and treatment side. For speech ability, the simple model included different speech outcomes,

hearing thresholds, and follow-up time points, while the adjusted model further included age group, treatment side, and the presence of cochlear implants before gene therapy. Results from the LMMs are presented as estimated means with 95% confidence intervals (CIs).

Distribution of treatment outcomes (efficacy and ineffacy) by baseline characteristics, including age group, the number of present DPOAE at baseline, and *OTOF* variants, was analyzed using Fisher's exact test. Comparisons of mean hearing thresholds at different time points were performed using the Kruskal–Wallis test followed by Dunn's multiple comparisons test. Improved thresholds represent the mean within-subject change from that individual's own baseline. Simple linear regression analysis was used to assess the relationship between the average SNR of DPOAE at baseline and the improvement in average ABR and ASSR thresholds at 26 weeks. Stata/MP (version 15.0) was used for statistical analyses, and a two-sided  $P < 0.05$  was considered to be statistically significant. GraphPad Prism 10 was used for graphing and statistical analyses. Adobe Illustrator 2025 was used for graphing.

## Supplementary Tables

**Supplementary Table 1 | Baseline characteristics and follow-up duration of participants by center after expansion to multicenter period<sup>a</sup>**

| Item/Center                                      | Center<br>1<br>N = 17 | Center<br>2<br>N = 3 | Center<br>3<br>N = 1 | Center<br>4<br>N = 1 | Center<br>5<br>N = 1 | Center<br>6<br>N = 1 | Center<br>7<br>N = 2 | Center<br>8<br>N = 1 | Total<br>N = 27 |
|--------------------------------------------------|-----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|-----------------|
| <b>Age at administration — yr</b>                |                       |                      |                      |                      |                      |                      |                      |                      |                 |
| Mean                                             | 9.4                   | 2.3                  | 1.9                  | 5.5                  | 2.3                  | 1.7                  | 2.7                  | 3.2                  | 6.9             |
| Range                                            | 0.8-32.3              | 1.3-3.6              | 1.9                  | 5.5                  | 2.3                  | 1.7                  | 2.7                  | 3.2                  | 0.8-32.3        |
| <b>Male sex —<br/>no. (%)</b>                    | 8 (47)                | 1 (33)               | 0                    | 1 (100)              | 1 (100)              | 1 (100)              | 1 (50)               | 0                    | 13 (48)         |
| <b>Hearing threshold<sup>b</sup></b>             |                       |                      |                      |                      |                      |                      |                      |                      |                 |
| ABR (dB<br>nHL)                                  | >95                   | >106                 | >100                 | >100                 | >100                 | >95                  | >100                 | >100                 | >97             |
| ASSR (dB<br>nHL)                                 | >98                   | >103                 | >98                  | >90                  | >93                  | 98                   | >90                  | >90                  | >97             |
| Behavioral<br>audiometry <sup>c</sup><br>(dB HL) | >97                   | >103                 | ND                   | 89                   | ND                   | 33                   | ND                   | ND                   | >95             |
| <b>OTOF variants — no. (%)<sup>d</sup></b>       |                       |                      |                      |                      |                      |                      |                      |                      |                 |
| NT/NT                                            | 4 (24)                | 0                    | 0                    | 1 (100)              | 0                    | 1 (100)              | 0                    | 0                    | 6 (22)          |
| NT/T                                             | 8 (47)                | 1 (33)               | 1 (100)              | 0                    | 1 (100)              | 0                    | 0                    | 0                    | 11 (41)         |
| T/T                                              | 5 (29)                | 2 (67)               | 0                    | 0                    | 0                    | 0                    | 2 (100)              | 1 (100)              | 10 (37)         |
| <b>Inclusion<br/>started</b>                     | Jul.<br>2024          | Aug.<br>2024         | Aug.<br>2024         | Sept.<br>2024        | Nov.<br>2024         | Jul.<br>2024         | Nov.<br>2024         | Apr.<br>2025         |                 |
| <b>Duration of<br/>follow-up</b>                 | W13-52                | W26-52               | W52                  | W26                  | W26                  | W26                  | W13                  | W13                  | W13-52          |

<sup>a</sup> Multicenter expansion began in July 2024 after the approval of Ethics Committee. In the period of multicenter expansion, 27 participants were enrolled across eight centers.

<sup>b</sup> Mean hearing threshold at 0.5, 1, 2, and 4 kHz; the mean hearing thresholds of all injected ears were included for baseline characterization. >, No response at maximum sound intensity level. ABR, auditory brainstem response. ASSR, auditory steady-state response.

<sup>c</sup> Only cooperative participants are included. ND, not done due to lack of cooperation.

<sup>d</sup> NT, non-truncating variants, including missense and in-frame variants. T, truncating variants, including nonsense, frameshift, large deletion, and splicing variants. NT/NT, biallelic non-truncating variants. NT/T, one allelic non-truncating variant and one allelic truncating variant. T/T, biallelic truncating variants.

**Supplementary Table 2 | Median (Range) for the cholesterol levels, lymphocyte counts, and other blood counts**

| <b>Type of Events</b>       | <b>1.5×10<sup>12</sup> vg/ear</b> | <b>2.1×10<sup>12</sup> vg/ear</b> |
|-----------------------------|-----------------------------------|-----------------------------------|
| Cholesterol high            | 5.82 (3.06) mmol/L                | 5.56 (1.83) mmol/L                |
| Lymphocyte counts increased | 7.30 (6.0)×10 <sup>9</sup> /L     | 7.98 (11.14)×10 <sup>9</sup> /L   |
| Neutrophil count decreased  | 1.09 (0.19)×10 <sup>9</sup> /L    | 1.18 (0.89)×10 <sup>9</sup> /L    |
| Platelet count decreased    | 183.5 (30.0)×10 <sup>9</sup> /L   | 159.0 (56.0)×10 <sup>9</sup> /L   |
| White blood cell decreased  | 4.29 (0.51)×10 <sup>9</sup> /L    | 4.98 (0.61)×10 <sup>9</sup> /L    |
| Eosinophilia                | NA                                | 0.74 (0.31)×10 <sup>9</sup> /L    |

NA, not applicable.

**Supplementary Table 3 | Sensitivity analysis of dose-dependent effects on ABR thresholds in participants with hearing recovery**

| ABR                   | Simple model               |           | Adjusted model             |           |
|-----------------------|----------------------------|-----------|----------------------------|-----------|
|                       | Estimated mean<br>(95% CI) | <i>P</i>  | Estimated mean<br>(95% CI) | <i>P</i>  |
| <b>Dose (vg/ear)</b>  |                            |           |                            |           |
| 1.5×10 <sup>12</sup>  |                            | Reference |                            | Reference |
| 2.1×10 <sup>12</sup>  | 0.30 (-5.90, 6.50)         | 0.93      | 1.29 (-8.57, 11.14)        | 0.80      |
| <b>Time</b>           |                            |           |                            |           |
| Pre                   |                            | Reference |                            | Reference |
| W6                    | -23.50 (-26.98, -20.02)    | <0.001    | -23.50 (-26.98, -20.02)    | <0.001    |
| W13                   | -30.29 (-33.77, -26.80)    | <0.001    | -30.29 (-33.77, -26.80)    | <0.001    |
| W26                   | -38.59 (-42.29, -34.89)    | <0.001    | -38.68 (-42.37, -34.98)    | <0.001    |
| W52                   | -40.96 (-45.05, -36.86)    | <0.001    | -41.07 (-45.16, -36.97)    | <0.001    |
| <b>Age (yr)</b>       |                            |           |                            |           |
| 0.5 to <3             |                            |           |                            | Reference |
| 3 to <8               |                            |           | -2.21 (-8.14, 3.73)        | 0.47      |
| 8 to <18              |                            |           | -13.55 (-28.17, 1.06)      | 0.07      |
| ≥18                   |                            |           | 8.62 (-4.00, 21.23)        | 0.18      |
| <b>Pre DPOAE</b>      |                            |           |                            |           |
| 1                     |                            |           |                            | Reference |
| 0                     |                            |           | -14.77 (-31.06, 1.52)      | 0.08      |
| 2                     |                            |           | -16.64 (-31.87, -1.42)     | 0.03      |
| 3                     |                            |           | -23.52 (-36.97, -10.06)    | 0.001     |
| 4                     |                            |           | -15.11 (-27.15, -3.08)     | 0.01      |
| 5                     |                            |           | -9.56 (-22.60, 3.48)       | 0.15      |
| <b>OTOF variants</b>  |                            |           |                            |           |
| NT/NT                 |                            |           |                            | Reference |
| NT/T                  |                            |           | -1.88 (-9.52, 5.76)        | 0.63      |
| T/T                   |                            |           | 0.32 (-7.16, 7.80)         | 0.93      |
| <b>Treatment side</b> |                            |           |                            |           |
| Unilateral            |                            |           |                            | Reference |
| Bilateral             |                            |           | 7.61 (-2.59, 17.81)        | 0.14      |

The simple model included average ABR thresholds, dose group, and follow-up time points (baseline to 52 weeks). The adjusted model further accounted for potential confounders, including age group, the number of present DPOAEs at baseline, *OTOF* variants, and treatment side. Pre DPOAE, the No. of present distortion product otoacoustic emissions (DPOAEs) at baseline. T/T, biallelic truncating variants. NT/T, one allelic non-truncating variant and one allelic truncating variant. NT/NT, biallelic non-truncating variants. ABR, auditory brainstem response.

**Supplementary Table 4 | Sensitivity analysis of dose-dependent effects on ASSR thresholds in participants with hearing recovery**

| ASSR                  | Simple model               |           | Adjusted model             |           |
|-----------------------|----------------------------|-----------|----------------------------|-----------|
|                       | Estimated mean<br>(95% CI) | <i>P</i>  | Estimated mean<br>(95% CI) | <i>P</i>  |
| <b>Dose (vg/ear)</b>  |                            |           |                            |           |
| 1.5×10 <sup>12</sup>  |                            | Reference |                            | Reference |
| 2.1×10 <sup>12</sup>  | 2.35 (-3.39, 8.10)         | 0.42      | -2.81 (-11.34, 5.72)       | 0.52      |
| <b>Time</b>           |                            |           |                            |           |
| Pre                   |                            | Reference |                            | Reference |
| W6                    | -24.26 (-28.22, -20.30)    | <0.001    | -24.26 (-28.22, -20.30)    | <0.001    |
| W13                   | -33.31 (-37.27, -29.35)    | <0.001    | -33.31 (-37.27, -29.35)    | <0.001    |
| W26                   | -39.38 (-43.58, -35.18)    | <0.001    | -39.36 (-43.55, -35.16)    | <0.001    |
| W52                   | -40.62 (-45.26, -35.97)    | <0.001    | -40.77 (-45.40, -36.13)    | <0.001    |
| <b>Age (yr)</b>       |                            |           |                            |           |
| 0.5 to <3             |                            |           |                            | Reference |
| 3 to <8               |                            |           | -6.60 (-11.75, -1.44)      | 0.01      |
| 8 to <18              |                            |           | -5.27 (-17.97, 7.43)       | 0.42      |
| ≥18                   |                            |           | 13.18 (2.07, 24.28)        | 0.02      |
| <b>Pre DPOAE</b>      |                            |           |                            |           |
| 1                     |                            |           |                            | Reference |
| 0                     |                            |           | -11.64 (-25.68, 2.41)      | 0.10      |
| 2                     |                            |           | -15.49 (-28.69, -2.30)     | 0.02      |
| 3                     |                            |           | -18.16 (-29.87, -6.46)     | 0.002     |
| 4                     |                            |           | -12.03 (-22.45, -1.60)     | 0.02      |
| 5                     |                            |           | -3.16 (-14.48, 8.17)       | 0.59      |
| <b>OTOF variants</b>  |                            |           |                            |           |
| NT/NT                 |                            |           |                            | Reference |
| NT/T                  |                            |           | 1.54 (-5.09, 8.17)         | 0.65      |
| T/T                   |                            |           | -2.35 (-8.88, 4.18)        | 0.48      |
| <b>Treatment side</b> |                            |           |                            |           |
| Unilateral            |                            |           |                            | Reference |
| Bilateral             |                            |           | -1.41 (-10.20, 7.37)       | 0.75      |

The simple model included average ASSR thresholds, dose group, and follow-up time points (baseline to 52 weeks). The adjusted model further accounted for potential confounders, including age group, the number of present DPOAEs at baseline, *OTOF* variants, and treatment side. Pre DPOAE, the No. of present distortion product otoacoustic emissions (DPOAEs) at baseline. T/T, biallelic truncating variants. NT/T, one allelic non-truncating variant and one allelic truncating variant. NT/NT, biallelic non-truncating variants. ASSR, auditory steady-state response.

**Supplementary Table 5 | Sensitivity analysis of dose-dependent effects on behavioral thresholds in participants with hearing recovery**

| Behavioral audiometry | Simple model            |           | Adjusted model          |           |
|-----------------------|-------------------------|-----------|-------------------------|-----------|
|                       | Estimated mean (95% CI) | <i>P</i>  | Estimated mean (95% CI) | <i>P</i>  |
| <b>Dose (vg/ear)</b>  |                         |           |                         |           |
| 1.5×10 <sup>12</sup>  |                         | Reference |                         | Reference |
| 2.1×10 <sup>12</sup>  | 1.87 (-5.18, 8.93)      | 0.60      | 1.14 (-9.50, 11.79)     | 0.83      |
| <b>Time</b>           |                         |           |                         |           |
| Pre                   |                         | Reference |                         | Reference |
| W6                    | -24.66 (-30.15, -19.16) | <0.001    | -24.51 (-30.03, -19.00) | <0.001    |
| W13                   | -35.74 (-41.07, -30.40) | <0.001    | -35.44 (-40.79, -30.09) | <0.001    |
| W26                   | -44.02 (-49.59, -38.45) | <0.001    | -43.95 (-49.54, -38.36) | <0.001    |
| W52                   | -48.92 (-55.05, -42.78) | <0.001    | -48.45 (-54.59, -42.31) | <0.001    |
| <b>Age (yr)</b>       |                         |           |                         |           |
| 0.5 to <3             |                         |           |                         | Reference |
| 3 to <8               |                         |           | 1.06 (-5.53, 7.65)      | 0.75      |
| 8 to <18              |                         |           | 10.16 (-8.92, 29.23)    | 0.30      |
| ≥18                   |                         |           | 3.43 (-8.89, 15.74)     | 0.59      |
| <b>Pre DPOAE</b>      |                         |           |                         |           |
| 1                     |                         |           |                         | Reference |
| 0                     |                         |           | -6.28 (-26.93, 14.38)   | 0.55      |
| 2                     |                         |           | -5.00 (-25.22, 15.23)   | 0.63      |
| 3                     |                         |           | -8.95 (-27.45, 9.55)    | 0.34      |
| 4                     |                         |           | -9.31 (-26.71, 8.08)    | 0.29      |
| 5                     |                         |           | -6.53 (-24.67, 11.61)   | 0.48      |
| <b>OTOF variants</b>  |                         |           |                         |           |
| NT/NT                 |                         |           |                         | Reference |
| NT/T                  |                         |           | 14.12 (5.97, 22.28)     | 0.001     |
| T/T                   |                         |           | 8.27 (0.41, 16.14)      | 0.04      |
| <b>Treatment side</b> |                         |           |                         |           |
| Unilateral            |                         |           |                         | Reference |
| Bilateral             |                         |           | -0.43 (-10.85, 9.99)    | 0.94      |

The simple model included average behavioral audiometry, dose group, and follow-up time points (baseline to 52 weeks). The adjusted model further accounted for potential confounders, including age group, the number of present DPOAEs at baseline, *OTOF* variants, and treatment side. Pre DPOAE, the No. of present distortion product otoacoustic emissions (DPOAEs) at baseline. T/T, biallelic truncating variants. NT/T, one allelic non-truncating variant and one allelic truncating variant. NT/NT, biallelic non-truncating variants.

**Supplementary Table 6 | Sensitivity analysis of the relationship between hearing thresholds and IT-MAIS/MAIS scores in participants with hearing recovery**

| IT-MAIS/MAIS                | Simple model               |           | Adjusted model             |           | IT-MAIS/MAIS                     | Simple model               |           | Adjusted model             |           |
|-----------------------------|----------------------------|-----------|----------------------------|-----------|----------------------------------|----------------------------|-----------|----------------------------|-----------|
|                             | Estimated mean<br>(95% CI) | <i>P</i>  | Estimated mean<br>(95% CI) | <i>P</i>  |                                  | Estimated mean<br>(95% CI) | <i>P</i>  | Estimated mean<br>(95% CI) | <i>P</i>  |
| <b>ABR</b>                  | -0.02 (-0.11, 0.07)        | 0.62      | -0.03 (-0.12, 0.05)        | 0.47      | <b>Behavioral<br/>audiometry</b> | -0.10 (-0.17, -0.03)       | 0.003     | -0.11 (-0.17, -0.04)       | 0.001     |
| <b>Time</b>                 |                            |           |                            |           | <b>Time</b>                      |                            |           |                            |           |
| Pre                         |                            | Reference |                            | Reference | Pre                              |                            | Reference |                            | Reference |
| W6                          | 6.87 (3.73, 10.02)         | <0.001    | 6.66 (3.57, 9.75)          | <0.001    | W6                               | 4.76 (1.74, 7.78)          | 0.002     | 4.74 (1.74, 7.75)          | 0.002     |
| W13                         | 13.41 (9.82, 17.00)        | <0.001    | 13.15 (9.62, 16.67)        | <0.001    | W13                              | 11.13 (7.73, 14.52)        | <0.001    | 11.04 (7.69, 14.39)        | <0.001    |
| W26                         | 20.23 (15.93, 24.53)       | <0.001    | 19.77 (15.57, 23.98)       | <0.001    | W26                              | 17.12 (13.14, 21.10)       | <0.001    | 16.93 (13.04, 20.83)       | <0.001    |
| W52                         | 26.91 (22.06, 31.77)       | <0.001    | 26.24 (21.46, 31.02)       | <0.001    | W52                              | 23.59 (19.10, 28.08)       | <0.001    | 23.22 (18.83, 27.61)       | <0.001    |
| <b>Age (yr)</b>             |                            |           |                            |           | <b>Age (yr)</b>                  |                            |           |                            |           |
| 0.5 to <3                   |                            |           |                            | Reference | 0.5 to <3                        |                            |           |                            | Reference |
| 3 to <8                     |                            |           | 0.70 (-3.02, 4.42)         | 0.71      | 3 to <8                          |                            |           | -0.43 (-3.38, 2.51)        | 0.77      |
| 8 to <18                    |                            |           | -2.28 (-7.79, 3.23)        | 0.42      | 8 to <18                         |                            |           | -1.46 (-5.87, 2.95)        | 0.52      |
| <b>Treatment side</b>       |                            |           |                            |           | <b>Treatment side</b>            |                            |           |                            |           |
| Unilateral                  |                            |           |                            | Reference | Unilateral                       |                            |           |                            | Reference |
| Bilateral                   |                            |           | 7.04 (3.19, 10.89)         | <0.001    | Bilateral                        |                            |           | 5.55 (2.55, 8.55)          | <0.001    |
| <b>Cochlear<br/>implant</b> |                            |           |                            |           | <b>Cochlear<br/>implant</b>      |                            |           |                            |           |
| Without<br>cochlear implant |                            |           |                            | Reference | Without<br>cochlear implant      |                            |           |                            | Reference |
| With<br>cochlear implant    |                            |           | 0.13 (-4.52, 4.79)         | 0.96      | With<br>cochlear implant         |                            |           | -0.74 (-4.47, 3.00)        | 0.70      |

The children's auditory abilities in real-life communication environment were primarily assessed using the Infant-Toddler Meaningful Auditory Integration Scale (IT-MAIS) and Meaningful Auditory Integration Scale (MAIS). The simple model included ABR/behavioral audiometry and follow-up time points (baseline to 52 weeks). The adjusted model further accounted for potential confounders, including age group, treatment side, and whether to receive a cochlear implant or not before gene therapy. ABR, auditory brainstem response.

**Supplementary Table 7 | Sensitivity analysis of the relationship between hearing thresholds and SSQP-noise related scores in participants with hearing recovery**

| SSQP-noise related       | Simple model            |           | Adjusted model          |           | SSQP-noise related           | Simple model            |           | Adjusted model          |           |
|--------------------------|-------------------------|-----------|-------------------------|-----------|------------------------------|-------------------------|-----------|-------------------------|-----------|
|                          | Estimated mean (95% CI) | P         | Estimated mean (95% CI) | P         |                              | Estimated mean (95% CI) | P         | Estimated mean (95% CI) | P         |
| <b>ABR</b>               | -0.04 (-0.08, -0.01)    | 0.005     | -0.05 (-0.08, -0.02)    | 0.001     | <b>Behavioral audiometry</b> | -0.05 (-0.07, -0.02)    | <0.001    | -0.05 (-0.07, -0.02)    | <0.001    |
| <b>Time</b>              |                         |           |                         |           | <b>Time</b>                  |                         |           |                         |           |
| Pre                      |                         | Reference |                         | Reference | Pre                          |                         | Reference |                         | Reference |
| W6                       | 0.85 (-0.26, 1.96)      | 0.13      | 0.71 (-0.39, 1.81)      | 0.21      | W6                           | 0.73 (-0.33, 1.80)      | 0.18      | 0.67 (-0.41, 1.74)      | 0.22      |
| W13                      | 2.24 (0.99, 3.49)       | <0.001    | 2.08 (0.83, 3.32)       | 0.001     | W13                          | 2.33 (1.15, 3.52)       | <0.001    | 2.29 (1.10, 3.49)       | <0.001    |
| W26                      | 3.92 (2.43, 5.41)       | <0.001    | 3.68 (2.20, 5.16)       | <0.001    | W26                          | 3.91 (2.54, 5.28)       | <0.001    | 3.84 (2.46, 5.22)       | <0.001    |
| W52                      | 4.69 (3.01, 6.37)       | <0.001    | 4.42 (2.74, 6.09)       | <0.001    | W52                          | 4.61 (3.05, 6.17)       | <0.001    | 4.57 (3.00, 6.14)       | <0.001    |
| <b>Age (yr)</b>          |                         |           |                         |           | <b>Age (yr)</b>              |                         |           |                         |           |
| 0.5 to <3                |                         |           |                         | Reference | 0.5 to <3                    |                         |           |                         | Reference |
| 3 to <8                  |                         |           | 0.38 (-0.99, 1.76)      | 0.58      | 3 to <8                      |                         |           | 0.27 (-1.03, 1.57)      | 0.69      |
| 8 to <18                 |                         |           | 0.00 (-2.04, 2.04)      | >0.99     | 8 to <18                     |                         |           | 0.35 (-1.58, 2.28)      | 0.72      |
| <b>Treatment side</b>    |                         |           |                         |           | <b>Treatment side</b>        |                         |           |                         |           |
| Unilateral               |                         |           |                         | Reference | Unilateral                   |                         |           |                         | Reference |
| Bilateral                |                         |           | 1.92 (0.49, 3.34)       | 0.008     | Bilateral                    |                         |           | 1.33 (0, 2.66)          | 0.05      |
| <b>Cochlear implant</b>  |                         |           |                         |           | <b>Cochlear implant</b>      |                         |           |                         |           |
| Without cochlear implant |                         |           |                         | Reference | Without cochlear implant     |                         |           |                         | Reference |
| With cochlear implant    |                         |           | -0.02 (-1.82, 1.77)     | 0.98      | With cochlear implant        |                         |           | -0.33 (-2.06, 1.40)     | 0.71      |

Speech perception in noise was assessed by the speech, spatial, and other qualities of hearing scale for parents (SSQ-P). We extracted some noise-related items from the SSQ-P, including items 1, 4, 5, 6, and 8 from the Speech section, and item 2 from the Other Qualities section. The simple model included ABR/behavioral audiometry and follow-up time points (baseline to 52 weeks). The adjusted model further accounted for potential confounders, including age group, treatment side and whether to receive a cochlear implant or not before gene therapy. ABR, auditory brainstem response.

**Supplementary Table 8 | Sensitivity analysis of the relationship between hearing thresholds and MUSS scores in participants with hearing recovery**

| MUSS                     | Simple model            |           | Adjusted model          |           | MUSS                         | Simple model            |           | Adjusted model          |           |
|--------------------------|-------------------------|-----------|-------------------------|-----------|------------------------------|-------------------------|-----------|-------------------------|-----------|
|                          | Estimated mean (95% CI) | P         | Estimated mean (95% CI) | P         |                              | Estimated mean (95% CI) | P         | Estimated mean (95% CI) | P         |
| <b>ABR</b>               | -0.01 (-0.10, 0.09)     | 0.86      | -0.01 (-0.10, 0.08)     | 0.90      | <b>Behavioral audiometry</b> | -0.13 (-0.21, -0.05)    | 0.001     | -0.14 (-0.22, -0.06)    | <0.001    |
| <b>Time</b>              |                         |           |                         |           | <b>Time</b>                  |                         |           |                         |           |
| Pre                      |                         | Reference |                         | Reference | Pre                          |                         | Reference |                         | Reference |
| W6                       | 1.19 (-2.26, 4.63)      | 0.50      | 1.26 (-2.10, 4.62)      | 0.46      | W6                           | -2.03 (-5.61, 1.55)     | 0.27      | -2.05 (-5.60, 1.49)     | 0.26      |
| W13                      | 3.52 (-0.40, 7.44)      | 0.08      | 3.57 (-0.23, 7.37)      | 0.07      | W13                          | -0.64 (-4.66, 3.37)     | 0.75      | -0.75 (-4.69, 3.20)     | 0.71      |
| W26                      | 11.68 (7.00, 16.35)     | <0.001    | 11.67 (7.16, 16.18)     | <0.001    | W26                          | 6.95 (2.26, 11.65)      | 0.004     | 6.57 (2.00, 11.15)      | 0.005     |
| W52                      | 21.80 (16.51, 27.10)    | <0.001    | 21.68 (16.53, 26.83)    | <0.001    | W52                          | 17.19 (11.88, 22.49)    | <0.001    | 16.73 (11.56, 21.89)    | <0.001    |
| <b>Age (yr)</b>          |                         |           |                         |           | <b>Age (yr)</b>              |                         |           |                         |           |
| 0.5 to <3                |                         |           |                         | Reference | 0.5 to <3                    |                         |           |                         | Reference |
| 3 to <8                  |                         |           | 1.06 (-2.45, 4.57)      | 0.55      | 3 to <8                      |                         |           | 0.73 (-2.64, 4.10)      | 0.67      |
| 8 to <18                 |                         |           | -1.23 (-6.45, 3.99)     | 0.64      | 8 to <18                     |                         |           | 0.62 (-4.43, 5.68)      | 0.81      |
| <b>Treatment side</b>    |                         |           |                         |           | <b>Treatment side</b>        |                         |           |                         |           |
| Unilateral               |                         |           |                         | Reference | Unilateral                   |                         |           |                         | Reference |
| Bilateral                |                         |           | 0.24 (-3.39, 3.87)      | 0.90      | Bilateral                    |                         |           | -0.91 (-4.34, 2.51)     | 0.60      |
| <b>Cochlear implant</b>  |                         |           |                         |           | <b>Cochlear implant</b>      |                         |           |                         |           |
| Without cochlear implant |                         |           |                         | Reference | Without cochlear implant     |                         |           |                         | Reference |
| With cochlear implant    |                         |           | 6.93 (2.52, 11.34)      | 0.002     | With cochlear implant        |                         |           | 6.46 (2.19, 10.73)      | 0.003     |

Meaningful Use of Speech Scale (MUSS) is used to assess the speech ability of patients with hearing loss. The simple model included ABR/behavioral audiometry and follow-up time points (baseline to 52 weeks). The adjusted model further accounted for potential confounders, including age group, treatment side and whether to receive a cochlear implant or not before gene therapy. ABR, auditory brainstem response.

## References

- 1 Sabo, D. L. The audiologic assessment of the young pediatric patient: the clinic. *Trends Amplif* **4**, 51-60 (1999). <https://doi.org/10.1177/108471389900400205>
- 2 Robbins, A. M., Renshaw, J. J. & Berry, S. W. Evaluating meaningful auditory integration in profoundly hearing-impaired children. *Am J Otol* **12 Suppl**, 144-150 (1991).
- 3 Robbins AM, O. M. *Meaningful Use of Speech Scale (MUSS)*. (Indiana University School of Medicine, 1990).
- 4 Pennini, P. T. M. & Almeida, K. Speech, Spatial and Qualities of Hearing Scale in assessing the benefit in hearing aid users. *Codas* **33**, e20190196 (2021). <https://doi.org/10.1590/2317-1782/20202019196>
- 5 Galvin, K. L. & Noble, W. Adaptation of the speech, spatial, and qualities of hearing scale for use with children, parents, and teachers. *Cochlear Implants Int* **14**, 135-141 (2013). <https://doi.org/10.1179/1754762812y.0000000014>
- 6 Archbold, S., Lutman, M. E. & Nikolopoulos, T. Categories of auditory performance: inter-user reliability. *Br J Audiol* **32**, 7-12 (1998). <https://doi.org/10.3109/03005364000000045>
- 7 McDaniel, D. M. & Cox, R. M. Evaluation of the speech intelligibility rating (SIR) test for hearing aid comparisons. *J Speech Hear Res* **35**, 686-693 (1992). <https://doi.org/10.1044/jshr.3503.686>
- 8 Fu, Q. J., Zhu, M. & Wang, X. Development and validation of the Mandarin speech perception test. *The Journal of the Acoustical Society of America* **129**, EL267-273 (2011). <https://doi.org/10.1121/1.3590739>
- 9 Cheng, X. *et al.* Music Training Can Improve Music and Speech Perception in Pediatric Mandarin-Speaking Cochlear Implant Users. *Trends in hearing* **22**, 2331216518759214 (2018). <https://doi.org/10.1177/2331216518759214>
- 10 Lv, J. *et al.* AAV1-hOTOF gene therapy for autosomal recessive deafness 9: a single-arm trial. *Lancet* **403**, 2317-2325 (2024). [https://doi.org/10.1016/s0140-6736\(23\)02874-x](https://doi.org/10.1016/s0140-6736(23)02874-x)
- 11 Wang, H. *et al.* Bilateral gene therapy in children with autosomal recessive deafness 9: single-arm trial results. *Nat Med* **30**, 1898-1904 (2024). <https://doi.org/10.1038/s41591-024-03023-5>