Indirect vs Direct Voice Therapy for Children With Vocal Nodules
A Randomized Clinical Trial

Christopher Hartnick, MD; Catherine Ballif, MA, CCC-SLP; Vanessa De Guzman, MS; Robert Sataloff, MD; Paolo Campisi, MD; Joseph Kerschner, MD; Adrianna Shembel, MA, CCC-SLP; Domenic Reda, PhD; Helen Shi, MS; Elinore Sheryka Zacny, MS; Glenn Bunting, MS, CCC-SLP

IMPORTANCE Benign vocal fold nodules affect 12% to 22% of the pediatric population, and 95% of otolaryngologists recommend voice therapy as treatment. However, no randomized clinical trials that we are aware of have shown its benefits.

OBJECTIVE To determine the impact of voice therapy in children with vocal fold nodules according to pretherapy and posttherapy scores on the Pediatric Voice-Related Quality of Life (PVRQOL) survey; secondary objectives included changes in phonatory parameters.

DESIGN, SETTING, AND PARTICIPANTS For this multicenter randomized clinical trial, 114 children ages 6 to 10 years with vocal fold nodules, PVRQOL scores less than 87.5, and dysphonia for longer than 12 weeks were recruited from outpatient voice and speech clinics. This age range was identified because these patients have not experienced pubertal changes of the larynx, tolerate stroboscopy, and cooperate with voice therapy. Participants were blinded to treatment arm.

INTERVENTIONS Participants received either indirect or direct therapy for 8 to 12 weeks. Indirect therapy focused on education and discussion of voice principles, while direct treatment used the stimulus, response, antecedent paradigm.

MAIN OUTCOMES AND MEASURES The primary outcome measure was PVRQOL score change before and after treatment. Secondary phonatory measures were also compared.

RESULTS Overall, 114 children were recruited for study (mean [SD] age, 8 [1.4] years; 83 males [73%]); with 57 randomized to receive either indirect or direct therapy. Both direct and indirect therapy approaches showed significant differences in PVRQOL scores pretherapy to posttherapy. The mean increase in PVRQOL score for direct therapy was 19.2, and 14.7 for indirect therapy (difference, 4.5; 95.3% CI, −10.8 to 19.8). Of 44 participants in the direct therapy group, 27 (61%) achieved a clinically meaningful PVRQOL improvement, compared with 26 of 49 (53%) for indirect therapy (difference, 8%; 95% CI, −12 to 28). Post hoc stratification showed robust effects in the direct therapy group for older children (Cohen $d = 0.50$) and the latter two-thirds of participants (Cohen $d = 0.46$). Vocal fold nodules reduced in size in 31% (22 of 70) and completely resolved in 11% (8 of 70) of participants who consented to a second set of images after going through the recruitment process.

CONCLUSIONS AND RELEVANCE Both direct and indirect voice therapy improved voice-related quality of life in children with vocal fold nodules, although there was no significant difference between approaches. Future studies may focus upon which voice therapy approaches are effective in treating age-defined populations.

TRIAL REGISTRATION clinicaltrials.gov Identifier: NCT01255735

Published online December 28, 2017.

© 2017 American Medical Association. All rights reserved.
Voice disorders affect over 5 million children,1,2 with at least 3 million children referred to speech-language pathologists (SLPs) for voice therapy because of persistent dysphonia3 caused by bilateral benign vocal fold nodules. These nodules present in 35% to 78% of cases of pediatric voice disorders,4,9 with higher rates in males (21.6%) than females (11.7%).4,10 Children who cannot express themselves adequately because of voice disorders may have underdeveloped communication skills and psychosocial abilities11 that are associated with poor self-esteem and self-consciousness.12 Studies also show children with unresolved voice disorders can require additional ongoing treatment into adulthood, placing substantial burden on the medical system.13

Although there have been recent studies focused on voice therapy, which 95% of otolaryngologists prescribe for pediatric vocal fold nodules,14-19 no study that we are aware of has determined its benefits. Our goal was to assess 2 treatment approaches: a direct treatment approach—Adventures in Voice (AIV)—and an indirect treatment approach—My Voice Adventure (MVA)—for children with vocal fold nodules. Our primary objective was to determine the effect of voice therapy on voice-related quality of life in children ages 6 to 10 years with vocal fold nodules as measured by change scores pre-therapy and post-therapy on the Pediatric Voice-Related Quality of Life (PVRQOL) survey19 and whether statistically significant differences exist in PVRQOL scores between these approaches. Our secondary objective was to determine whether acoustic, aerodynamic, auditory-perceptual, and anatomical parameters changed with treatment and whether these secondary outcomes demonstrated better improvement in a single specific approach.

Methods

Participants

One hundred fourteen (114) participants (83 boys; 31 girls) between the ages of 6 and 10 years were recruited at 4 voice clinics over 3 years (Table 1). This age group was chosen because these children (1) have not yet undergone pubertal changes that affect the larynx, thereby ensuring homogeneity; (2) can tolerate stroboscopy; and (3) can cooperate with voice therapy (Figure). Power analysis calculations projected a total sample size of 92, based on the mean (SD) difference in PVRQOL change scores of 8 (11.7), with independent sample t tests, 2-sided α of 0.05 and power of 0.90. The sample size was adjusted to account for a withdrawal rate of 20%, bringing the final sample size to 116 (n = 58 AIV; n = 58 MVA). Additionally, because vocal fold nodules affect male children more than female children (7:3 ratio), projected enrollment included 70% males and 30% females.4,10 Study statisticians generated permuted block randomization schemes using SAS 9.2 (SAS Institute) and stratified by study site, age (6-7 years old vs 8-10 years old), and sex. Once a participant met eligibility criteria (eTable 1 in the Supplement), each site’s study coordinator assigned the participant to either AIV or MVA using a web-based randomization system. All participants were blinded to study arm. Sex, race, and ethnicity were collected as required by the study sponsor.

Findings

In this randomized clinical trial, both therapy approaches showed significant pre-therapy to post-therapy improvement in Pediatric Voice-Related Quality of Life Survey scores. No difference was observed between approaches.

Meaning

Both direct and indirect voice therapy approaches may have a role in the treatment of pediatric vocal fold nodules, particularly in treating defined-age populations.

Trial protocol is available in the Supplement.

Procedures

The study occurred at Massachusetts Eye and Ear, Medical College of Wisconsin, Drexel University College of Medicine, and Hospital for Sick Children. Each site’s institutional review board approved the study. Each study had at least 2 SLPs: a treating SLP who administered treatment and a blinded, nontreating SLP who obtained secondary outcome data.

All treating SLPs were trained in both treatment protocols. Clinicians video-recorded practice sessions that were reviewed by the training SLP, who provided feedback until therapeutic skills were satisfactory to treat study participants. To ensure treatment consistency across sites, clinicians met monthly to review each protocol and address specific questions. All nontreating SLPs used the KeyPENTAX Phonatory Aerodynamic System Model 6600 and Computerized Speech Laboratory (PENTAX Medical) according to their respective protocols to obtain secondary outcome data such as Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) sentences, phonation threshold pressures, and noise-to-harmonics ratio. Furthermore, 2 independent and blinded raters were selected to rate the CAPE-V sentences. Training was not provided because both raters had extensive experience with the CAPE-V.

Each participant underwent stroboscopy, and the presence of bilateral vocal fold nodules was confirmed by all investigators. Once participants or their parents or caregivers provided written informed consent and met all eligibility criteria, a web-based randomization system assigned them to a treatment arm, and they were then scheduled for voice therapy with the treating SLP.

Key Points

- **Question**: What is the effect of indirect vs direct voice therapy on pediatric patients with vocal nodules?
- **Findings**: In this randomized clinical trial, both therapy approaches showed significant pre-therapy to post-therapy improvement in Pediatric Voice-Related Quality of Life Survey scores. No difference was observed between approaches.
- **Meaning**: Both direct and indirect voice therapy approaches may have a role in the treatment of pediatric vocal fold nodules, particularly in treating defined-age populations.

- **Trial protocol is available in the Supplement.**

- **Procedures**: The study occurred at Massachusetts Eye and Ear, Medical College of Wisconsin, Drexel University College of Medicine, and Hospital for Sick Children. Each site’s institutional review board approved the study. Each study had at least 2 SLPs: a treating SLP who administered treatment and a blinded, nontreating SLP who obtained secondary outcome data.

- **Intervention**: For the purposes of this study, the terms “direct” and “indirect” differentiate treatment arms. It should not be assumed that “indirect” treatment did not contain any elements of direct therapy—both therapy arms contained direct therapy techniques. However, indirect treatment focused more on education/discussion of voice principles, whereas direct treatment used the stimulus, response, antecedent paradigm and incorporated more shaping and practice.

- **The direct therapy arm, AIV—created by Katherine Verdolini-Abbott, PhD, CCC-SLP, an independent consultant to the...**
Establishing new voice patterns through motor learning principles (AIV) involved techniques such as: (1) identifying phonotraumatic behaviors, as well as environments that result in dysphonia; (2) practice of fundamentals twice daily (each with its own compliance data point); and (3) a specific treatment involved games and other activities to increase awareness and change phonotraumatic vocal behaviors. Specific areas of focus included:

- Basic education of the vocal mechanism, normal function, and general care.
- Identifying phonotraumatic behaviors, as well as environments and/or situations and ways to avoid or modify behaviors.
- Discussion of “undesired” and “desired” voice quality and production.

While both treatment approaches appear similar because they bring attention to new vocal productions, focus on differences between “old” and “new” phonatory patterns, and stress the importance of therapeutic compliance, they differ in how treatment is delivered. My Voice Adventure discusses ways to modify or eliminate behavior within an environment while AIV changes vocal production via direct phonatory modification, somatosensory input, discrimination, and practice. Despite these differences, however, the treatment arms were developed to mirror each other in terms of timeframe (8-12 weeks) and prize incentives participants received for compliance.

Treatment compliance was based on how many of the homework assignments the participant completed rather than the quality of the homework. For example, the MVA protocol had 3 daily compliance points. The participant reported every day on 2 areas: rating of voice function for that day on a scale of 0 to 10, where 10 was the best possible voice and the total number of glasses of fluid consumed that day. If the participant reported on both areas, they received an additional compliance point. Similarly, in the AIV arm, participants had 3 to 4 homework daily assignments to complete based on their individualized plan. All AIV participants reported on: (1) vocal hygiene practices; (2) practice of fundamentals twice daily (each with its own compliance data point); and (3) a specific assignment or activity targeting the module(s) addressed in that week’s session. In both protocols, the participant was rewarded with small age-appropriate toys for completing and returning their daily data sheets rather than for reporting completion of homework on a daily basis or for the quality of their practice. No adverse events were reported.
<table>
<thead>
<tr>
<th>Measure</th>
<th>Mean (95% CI) [Participants No.]</th>
<th>AIV</th>
<th>MVA</th>
<th>Difference</th>
<th>Cohen d&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>PVRQOL Survey score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>64.0 (59.2 to 68.8) [n = 56]</td>
<td>71.1 (70.0 to 75.2) [n = 56]</td>
<td>7.1 (0.8 to 13.3) ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intention-to-treat post therapy change (visit 3)</td>
<td>19.2 (−19.4 to 57.9) [n = 56]</td>
<td>14.7 (−13.8 to 43.3) [n = 56]</td>
<td>4.5 (−10.5 to 19.5) ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAPE-V Sentences auditory perceptual rating score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>32.5 (29.6 to 35.3) [n = 51]</td>
<td>33.4 (30.5 to 36.2) [n = 51]</td>
<td>−0.9 (−4.9 to 3.1) ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post therapy change (visit 3)</td>
<td>−6.2 (−9.5 to −3.0) [n = 42]</td>
<td>−6.7 (−10.4 to −3.0) [n = 41]</td>
<td>−0.5 (−5.3 to 4.4) ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Noise-to-harmonic ratio score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>0.113 (0.110 to 0.119) [n = 53]</td>
<td>0.115 (0.110 to 0.120) [n = 49]</td>
<td>0.002 (−0.006 to 0.010) ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post therapy change (visit 3)</td>
<td>−0.007 (−0.021 to 0.007) [n = 42]</td>
<td>0.005 (−0.001 to 0.011) [n = 42]</td>
<td>0.012 (−0.003 to 0.027) ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phonation threshold pressures score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>7.6 (6.4 to 8.8) [n = 50]</td>
<td>7.3 (5.9 to 8.7) [n = 49]</td>
<td>−0.3 (−2.1 to 1.5) ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post therapy change&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.6 (−0.8 to 1.9) [n = 38]</td>
<td>0.5 (−0.6 to 1.7) [n = 40]</td>
<td>0.0 (−1.8 to 1.7) ND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PVRQOL Survey score change by age group, y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-7</td>
<td>18.1 (9.6 to 26.6) [n = 26]</td>
<td>17.1 (−11.0 to 23.1) [n = 26]</td>
<td>1.0 (−11.3 to 9.2) 0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8-10</td>
<td>23.0 (13.3 to 32.6) [n = 18]</td>
<td>13.6 (7.6 to 19.5) [n = 23]</td>
<td>9.4 (−1.7 to 20.4) 0.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PVRQOL Survey score change by sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>23.9 (16.8 to 31.0) [n = 33]</td>
<td>16.8 (11.8 to 21.9) [n = 36]</td>
<td>7.0 (−1.4 to 15.5) 0.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>8.8 (−3.0 to 20.5) [n = 11]</td>
<td>11.6 (4.0 to 19.2) [n = 13]</td>
<td>−2.8 (−15.6 to 10.0) 0.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PVRQOL Survey score change by site</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEEL, Massachusetts Eye and Ear Infirmary (Boston, Massachusetts)</td>
<td>27.7 (18.3 to 37.1) [n = 21]</td>
<td>22.9 (16.9 to 29.0) [n = 20]</td>
<td>4.8 (−6.1 to 15.7) 0.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical College of Wisconsin (Milwaukee, Wisconsin)</td>
<td>10.3 (3.8 to 16.7) [n = 12]</td>
<td>13.4 (5.8 to 21.1) [n = 16]</td>
<td>−3.2 (−13.2 to 6.9) 0.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hospital for Sick Children (Toronto, Ontario, Canada)</td>
<td>16.3 (6.0 to 32.1) [n = 11]</td>
<td>6.4 (−0.1 to 13.0) [n = 13]</td>
<td>9.9 (−6.6 to 26.4) 0.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PVRQOL Survey score change by enrollment period</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Earlier one-third enrolled</td>
<td>13.8 (5.3 to 22.3) [n = 15]</td>
<td>19.5 (12.1 to 27.0) [n = 17]</td>
<td>−5.7 (−16.5 to 5.0) 0.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latter two-thirds enrolled</td>
<td>23.4 (15.0 to 37.8) [n = 29]</td>
<td>13.3 (8.2 to 18.3) [n = 32]</td>
<td>10.1 (0.5 to 19.7) 0.46</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations. AIV, Adventures in Voice (direct vocal therapy); CAPE-V, Consensus Auditory-Perceptual Evaluation of Voice; MVA, My Voice Adventure (indirect vocal therapy); PVRQOL, Pediatric Voice-Related Quality of Life.<sup>a</sup> Change in posttherapy scores were measured after visit 3. <sup>b</sup> Interpretation of the Cohen d value is as follows: small effect = 0.2; medium effect = 0.5; large effect = 0.8.
Outcome Measures
The caregivers of study participants completed the PVRQOL survey at: (1) screening, immediately prior to randomization (baseline or visit 1); (2) the posttherapy evaluation conducted immediately following completion of therapy (visit 2); and (3) 1 month posttherapy (visit 3). The primary outcome measure was PVRQOL change scores from baseline to visit 3 to assess generalization and treatment carryover. Post hoc analysis was also performed on change scores at baseline and visit 2, as well as between baseline and visit 3. The PVRQOL contains 10 questions completed by the caregiver that takes approximately 5 minutes to complete (eFigure 1 in the Supplement). The scale ranges from 0 (worst) to 100 (best). It was previously validated in the pediatric population with voice disorders and has shown to have robust test-retest reliability and discriminant validity. A change score of 12 points is clinically significant.

The following secondary outcome measures were acquired at visit 1, visit 2, and visit 3: auditory-perceptual measures of vocal quality with CAPE-V sentences and phonation threshold pressures, noise-to-harmonics ratio, and nodule size and contour. Nodule size and contour was determined by a validated grading system with proven interrater reliability created by Shah et al, which categorizes nodules by their size (eg, grade 1, a small nodule; grade 2, a moderate nodule; or grade 3, a large nodule) and contour (discrete or sessile). Consideration was given to making nodule grade the primary outcome. However, data show that nodule size reduction or resolution do not strictly correlate with voice improvements. In fact, patients may have resolution to normal voice, even in the presence of residual nodules. Furthermore, nodule size may not correspond to the level of vocal dysfunction. Thus, although it is a concrete physical measurement, it was more appropriate as a secondary outcome, rather than a primary outcome.

Statistical Analysis
Statistical analysis was conducted using intention-to-treat principles, wherein all randomized participants were included in the primary outcome analysis. Multiple imputation was used for 19 participants with missing final PVRQOL scores (AIV, 12 of 56; MVA, 7 of 56). Sensitivity analysis was performed for missing data in 2 extreme situations. One set of analysis concentrated only on participants who completed the study, and another set concentrated on all noncompleters as no PVRQOL score change from baseline. t Tests and the model adjustments to the baseline PVRQOL run in the primary analysis were also performed in the sensitivity analysis; the results are consistent with the primary analysis that showed there is no evidence to show that assumption for Multiple Imputations—missing at random—is violated. All missing values were imputed and included in the primary outcome analysis. Regression analysis was performed to model PVRQOL change scores by treatment while controlling for baseline PVRQOL scores. To adjust for a potential ceiling effect associated with lesser potential improvement gains for participants with higher baseline PVRQOL measures, a second model was run using adjusted change score as a percentage of total possible improvement (defined as the magnitude of change from baseline to visit 3, divided by the difference between 100 and baseline score). Regression models were used to perform analysis on treatment differences for phonation threshold pressures, CAPE-V ratings, and noise-to-harmonics ratio change by controlling baseline score, age, and sex. Concordance correlation coefficients were run to determine interrater agreements between the CAPE-V raters. Effect size was reported as the difference in values and Cohen d. The precision of the effect size is presented with 95% CIs. The Cohen d was interpreted as small effect = 0.2; medium effect = 0.5; large effect = 0.8.

Results
Demographics
Of the 114 randomized participants (57 AIV; 57 MVA) in the study, 93 participants (46 AIV; 47 MVA) completed the study. Nineteen participants (10 AIV; 9 MVA) were lost to attrition. Mean compliance rates showed compliance in 85.4% of participants in the AIV group and in 91.2% of participants in the MVA group. Two participants (1 in each treatment group) were randomized because of clerical error and were excluded from statistical analysis. Table 1 provides participant demographics.

Primary Outcome: PVRQOL Survey Scores
Baseline PVRQOL raw scores are presented in Table 2 and differences between groups, while statistically significant, are not likely to be clinically meaningful. The original analysis plan compared the mean changes between groups. However, because of the baseline imbalance, we analyzed the data several ways: (1) as originally planned; (2) adjusting for the baseline score as a covariate; (3) and calculating the changes scores as score change/(100 − baseline score) to account for these baseline differences in the groups. Since results of the 3 analyses agreed, we decided to show only the latter analysis. There were clinical improvements in the intention-to-treat analysis of PVRQOL change scores from baseline to posttreatment in both AIV and MVA groups, with mean change score increases of 19.2 (95% CI, −19.4 to 57.9) and 14.7 (95% CI, −13.8 to 43.3), with a mean change score difference of 4.5 (95.3% CI, −10.8 to 19.8) between them. Sensitivity analysis including only completers showed a 20.1 (95% CI, 13.9 to 26.3) improvement in PVRQOL for the AIV group and 15.4 (95% CI, 11.3 to 19.8) in the MVA group. Furthermore, of the 44 participants in the AIV group, 27 (61%) achieved a clinically meaningful improvement in PVRQOL scores, compared with 26 of 49 (53%) in the MVA group for a difference of 8 percentage points (95% CI, −12 to 28).

Secondary Outcomes
CAPE-V Sentences
CAPE-V auditory-perceptual ratings, rated on a 0 to 100 mm visual analog scale, were similar across the groups at baseline (difference, −0.9; 95% CI, −4.9 to 3.1). On average, CAPE-V overall severity ratings decreased by 6.2 points in the AIV group and 6.7 points in the MVA group after treatment (difference,
0.5; 95% CI, −5.3 to 4.4) (Table 2). Agreement between raters for the overall severity CAPE-V subscore was evaluated by concordance correlation coefficient and found to be 0.63 (95% CI, 0.61 to 0.64) for overall severity change from baseline to visit 3.

### Noise-to-Harmonics Ratios

Baseline noise-to-harmonics ratios were similar between groups (difference, 0.002; 95% CI, −0.006 to 0.010). Noise-to-harmonics ratios decreased from baseline to visit 3 to 0.007 (95% CI, −0.021 to 0.007) in the AIV group and increased by 0.005 (95% CI, −0.001 to 0.011) in the MVA group; the difference in change scores between groups was 0.012 (95% CI, −0.003 to 0.027) (Table 2).

### Phonation Threshold Pressures

Mean phonation threshold pressures at baseline were similar across the groups (difference, −0.3; 95% CI, −2.1 to 1.5). Phonation threshold pressures increased by 0.6 in the AIV group and 0.5 in the MVA group (difference, 0.0; 95% CI, −1.8 to 1.7) (Table 2).

### Nodule Ratings

For those who improved in either group (ie, had reduced nodule size or no presence of vocal nodules), 31.4% (22 of 70) saw diminished nodules, and 4 participants from each group (n = 8 [11%]) had no vocal nodules posttreatment. For those who improved, nodule size improved more in the AIV group (n = 14 of 36 [39%]) vs the MVA group (n = 8 of 34 [24%]); difference 15% (95% CI, −6% to 37%). Four participants had worsened nodules: 3 participants in the AIV group and 1 participant in the MVA group.

### Compliance

Compliance was reported as a variable of percent compliance on a scale of 0 to 100. It was calculated by the number of homework completed divided by the number of homework assigned. There was a significant difference in compliance between the 2 treatment arms; however, in longitudinal models of PVRQOL scores and change scores over time (visit), treatment group, and compliance scores, compliance did not have a significant relationship with outcomes, and so compliance was dropped as a covariate from inclusion in additional models.

### Post Hoc Analysis: PVRQOL

Secondary post hoc analysis for PVRQOL change scores was conducted regarding age (Table 2). While the results of the adjusted analysis also showed no statistically significant differences between groups, a clinically significant effect was seen. Specifically, the Cohen d showed a medium effect size between treatment approaches in older children (8-10 years), favoring the AIV group (Cohen d = 0.50). Additional post hoc analysis stratified by enrollment period (Table 2) showed in the latter two-thirds of participants enrolled a difference between treatment arms in change score of 10.1 (95% CI, 0.5 to 19.7) favoring the AIV group (Cohen d = 0.46), suggesting a possible learning curve for administration of the AIV therapy.

### Discussion

Both direct and indirect voice therapy approaches improve voice-related quality of life in children ages 6 to 10 years with vocal fold nodules although neither significantly showed a difference to each other. It must be emphasized that there was no true placebo control arm in this study because our data and safety monitoring board (DSMB) deemed it unethical given the state of equipoise such that the improvement seen by both forms of voice therapy could be explained by therapy in general—an expectation by caregivers that the time was worth the investment or regression to the mean. While general benefit from voice therapy is one possible explanation for improvement, it may also be owing to whether the PVRQOL is sensitive enough to show a difference between the 2 groups when both groups show improvement. Previous studies—the majority conducted in adult participants—show more substantial improvements in direct therapy approaches compared with indirect approaches and control arms. However, a priori analysis in this study did not find any statistically significant differences in direct and indirect therapy approaches between the 2 groups, This incongruence may be related to the population in question (adults) and the cause of these differences (behavioral, developmental, cognitive, or other) warrants further investigation. We will consider other possible control group studies in the future.

### Limitations

Interestingly, although statistically significant and clinically meaningful improvements were seen in PVRQOL scores pretherapy to posttherapy, these improvements were not reflected in auditory-perceptual, acoustic, or aerodynamic outcomes owing to a few possibilities. First, there may be flaws in the methods themselves, such as the CAPE-V 0- to 100-mm visual analog scale ratings used to assess overall severity of aberrant voice quality. Auditory-perceptual impressions have shown to have poor-to-moderate interrater reliability, as internal standards can vary between raters, and numerical representations of perceptual measurements can be inherently unstable. Patient presentation (order effect) in the clinical setting has also shown to create additional instability in auditory-perceptual ratings. The moderate inter-rater agreements (concordance correlation coefficient = 0.63) found in this study, combined with high variability in the ratings themselves, is consistent with these shortcomings. Inherent instability in quantifying perceptual constructs of voice quality could explain why there was no auditory-perceptual difference found between pretherapy and posttherapy, despite participant-perceived improvements with therapy. Additionally, the fact that improved results reflected in PVRQOL change scores mirrored improvements in ratings of vocal quality by blinded, independent raters further illustrates methodological concerns with CAPE-V ratings.

Conversely, lack of significant findings in secondary outcomes, despite positive voice-related quality-of-life findings, could have less to do with methodology and more to do with psychosocial factors underlying behavioral interventions. Spe-
cifically, psychotherapeutic literature shows that interactions and interpersonal relationships between the therapist and patient improve the patient's perception of their condition. However, while pediatric voice therapy studies with control groups are considerably lacking a placebo group (mainly owing to ethical considerations from withholding treatment from children), previous adult studies show improved outcomes have more to do with the actual treatment than with patient-clinician rapport.

Furthermore, post-hoc analyses suggest that age may influence how well children respond with specific therapeutic approaches and should be considered when individualizing treatment. The robust clinical effect seen with older children (8-10 years) in the direct therapy group could be related to developmental differences between children ages 8 to 10 and 6 to 7 years. Hypothetically, the older the child, the better the child's cognitive development, giving older children an advantage when learning new vocal patterns, which is the hallmark of the AIV direct therapy approach. Further studies in this area should be investigated.

Finally, the last limitation may be the imbalance in raw PVRQOL scores between the 2 therapy groups at baseline. Notably, the study design initially required 2 separate administrations of baseline PVRQOL scores: (1) during study screening (conducted at the initial consultation) and (2) immediately prior to randomization (right before the start of therapy). The baseline scores were administered twice to account for prolonged periods between the initial clinical consultation and the first therapy session as an artifact of busy voice clinics involved in the study. Although the initial administration of the PVRQOL at screening showed no differences between randomized groups, the scores at screening could not be used due to, at times, substantial delays in the start of treatment. Our DSMB concluded that since the imbalance was not related to knowledge of the specific treatment assignment, study adjustments to the randomization process were unnecessary. Furthermore, the study statisticians and DSMB performed an interim analysis and found no evidence of problems with the randomization process. Finally, post-hoc analysis on PVRQOL change scores from screening to visit 3 also showed no significant differences between the 2 treatment groups, just as in the change scores at pre-randomization and visit 3.

Conclusions

Both direct and indirect voice therapy approaches appear to improve voice-related quality of life in children ages 6 to 10 years with vocal fold nodules although neither significantly showed a difference as opposed to the other. Although the improvement seen by both forms of voice therapy may be explained owing to parent expectations that the time was worth the investment or regression to the mean, it suggests that such improvement may be explained due to therapy in general. We will consider other possible control group studies, as well as studies that focus on which vocal therapy approaches are more effective in treating defined age populations.


