Voice Abnormalities at School Age in Children Born Extremely Preterm

WHAT'S KNOWN ON THIS SUBJECT: Isolated case reports of abnormal voice after extremely preterm birth are well described; however, there are no systematic studies of long-term voice outcomes in children born preterm.

WHAT THIS STUDY ADDS: Significant voice abnormalities were found in more than half of tested children born before 25 weeks' gestation. Multivariable analyses showed that the number of intubations, not the duration of intubation, and female gender were strongly associated with this adverse outcome.

BACKGROUND AND OBJECTIVES: Voice abnormality is a frequent finding in school age children born at <25 weeks' gestation in Western Australia. The objective of this study was to determine the frequency of voice abnormality, voice-related quality of life, and demographic and intubation factors in this population.

METHODS: Survivors <25 weeks' gestational age in Western Australia born from 1996 to 2004 were included. Voice assessments (auditory perceptual assessment scale and Pediatric Voice Handicap Index) were carried out by speech pathologists. Intubation history was obtained by retrospective chart review.

RESULTS: Of 251 NICU admissions, 154 (61%) survived. Exclusions were based on severe disability (11) or distant residence (13). Of 70 assessed, 67 completed assessments, 4 (6%) were in the normal range and 39 (58%) showed moderate-severe hoarseness. Simultaneous modeling of demographic and intubation characteristics showed an increased odds of moderate-severe voice disorder for children who had more than 5 intubations (odds ratio 6.96, 95% confidence interval 2.07–23.40, P = .002) and for girls relative to boys (odds ratio 3.46, 95% confidence interval 1.12–10.62, P = .030). Tube size and duration of intubation were not significant in the multivariable model. Median scores of parent-reported voice quality of life on the Pediatric Voice Handicap Index were markedly different for preterm (22) and term (3) groups, P < .001.

CONCLUSIONS: Voice disorders in this population were much more frequent than expected. Further studies are required to assess voice across a broader range of gestational ages, and to investigate voice-protective strategies in infants requiring multiple episodes of intubation.
Extreme preterm birth (<27 weeks’ gestation) is associated with adverse medical and social outcomes. Morbidity has been shown to be inversely related to gestational age. Long-term outcome studies to school age and beyond have identified a number of conditions more prevalent in the surviving population of extremely preterm infants. Voice quality has not been included in these long-term outcomes and reported only in isolated case reports.

Extremely preterm infants usually require respiratory support, typically endotracheal intubation, in the neonatal period because of physiologic immaturity. Dysphonia is a recognized complication of endotracheal intubation. Laryngeal injuries visualized postextubation in neonatal subjects range from mild erythema, edema, and granulation of the vocal folds to arytenoid cartilage dislocation, subglottic stenosis, laryngeal tears, and avulsion of the vocal folds. Surgical ligation of a patent ductus arteriosus (PDA) has been frequently associated with left vocal fold paralysis in some series. Several authors have identified a relationship between intubation factors and neonatal laryngeal injury. It remains unclear which factors, such as length of intubation, frequency of reintubation, and size of tube, are directly linked to persistent laryngeal disorders affecting voice quality.

In our long-term follow-up program, voice abnormality has been a common finding, particularly in children born at <25 weeks’ gestation. Voice outcomes in studies of children intubated in the neonatal period have so far been limited to those undergoing laryngeal reconstruction for recognized airway complications or after surgical ligation of the PDA, and to extremely low birth weight (ELBW) infants at 12 months of corrected age. Mild voice abnormalities, such as breathiness and roughness, are common in young children because of vocal overuse in childhood. Many of these cases will self-resolve, but more severe abnormalities require further investigation and intervention. Voice assessments typically involve auditory perceptual measures, quality-of-life assessments, and acoustic analysis, although the latter is rarely reported in children and has not been validated in the pediatric population. It is apparent that children as young as 5 are capable of reflecting on the social, emotional, and physical aspects of their voice abnormalities, and consequent negative emotional experiences have been reported. Among other things, children with chronic dysphonia cite concerns with participation limitation, negative evaluation of their voice quality by others in their social and academic environments, and managing emotions, such as frustration. Therefore, dysphonia has significant, negative effects on the quality of life of affected children. Little is known about the nature and incidence of later voice disorders in preterm children who were intubated after birth. In this study, we aimed to further explore demographic and intubation factors associated with voice abnormality in this population at school age and to assess the voice-related quality of life of these children.

METHODS

Patients were recruited from the neonatal follow-up program of the single tertiary neonatal service, which provides neonatal intensive care to all extremely preterm infants born in Western Australia, across 2 sites at King Edward Memorial Hospital and at Princess Margaret Hospital for Children. Cases selected were <25-week survivors between 6 and 15 years of age, born from 1996 to 2004, inclusive. Age criteria were selected to optimize compliance with voice assessments, which required reading a sample of connected speech and producing prolonged vowel sounds. Data from the neonatal follow-up program were used to exclude cases in which known disability was likely to prevent adequate assessment, or in which the family lived more than 200 km from the study center.

Neonatal clinical variables were available from the NICU database, and included demographic variables, clinical morbidities, duration and type of respiratory support, and duration of oxygen therapy. In addition, chart review was undertaken to determine additional information relating to each intubation/reintubation episode, such as type of endotracheal tube (ETT), size of ETT, route (oral or nasal), and body weight at the time of each intubation episode. A ratio of ETT size relative to body weight at the time of intubation/reintubation was created by dividing ETT size in millimeters by body weight in kilograms on the day of the procedure. The maximum ETT ratio of all intubation episodes was recorded for each infant.

Throughout the study period, uncuffed siliconized PVC (Portex, internal diameter 2.0–3.5 mm) or uncuffed ivory PVC (Portex, internal diameter 2.5–3.5 mm) ETTs were used. (Smiths Medical Australia Pty Ltd, Brisbane, Australia) Eight percent of all intubation episodes were with ivory tubes. Before 2000, ETT placement was predominantly nasotracheal after initial orotracheal placement for labor ward resuscitation, and after 2000, placement was predominantly orotracheal for all episodes. Elective versus emergency reintubation or degree of difficulty of intubation were insufficiently documented in the chart to allow analysis of these variables. Surgical ligation of a PDA is rarely carried out in extremely preterm infants in Western Australia, but was recorded where this occurred. Premedication before intubation was not used in this population.

Voice Assessments

1. The GRBAS is a widely used auditory perceptual assessment scale that determines ratings for each
of 5 voice characteristics: G (grade or severity of hoarseness), R (roughness: abnormal variations in pitch and loudness), B (breathiness: occurs when there is excess air heard in the voice), A (asthenicity: weakness of the voice), and S (strain: excess muscle tension used to produce voice). Each parameter is scored by using a 4-point ordinal scale from 0 to 3 (0, normal; 1, mild; 2, moderate; and 3, severe), and the mean of the values summed as the mean GRBAS. The G severity (hoarseness, 0–3 as above) was used as the primary outcome measure.

2. The Pediatric Voice Handicap Index (pVHI) is a quality-of-life questionnaire completed by parents and measures the impact of a voice disorder.51 The pVHI is scored on a 5-point Likert scale presented in 3 subscales each of 7 to 9 questions: (1) functional (eg, “At home we have difficulty hearing my child....”), (2) physical (eg, “My child uses a great deal of effort to speak....”), and (3) emotional (eg, “My child is frustrated with his/her voice problem....”). The maximum score attainable from these questions is 92, whereas normophonic control populations typically have total scores ≤2.

A term-born comparison group of 40 children was also recruited and they underwent the same measures. These children were recruited with parental consent from an orthopedic clinic of the children’s hospital, and all were born at term with no history of previous intubation or recent respiratory illness. Although not matched with the preterm group, they were of similar age and gender distribution.

The GRBAS was administered by 2 speech pathologists with postgraduate experience in clinical voice assessment. Inter- and intrarater reliability were calculated on a random sample of 10% of the raw voice samples.

**Statistical Method**

Continuous data were summarized by using median, interquartile range, and categorical data were summarized by using frequency distributions. The interrater and intrarater reliability for the GRBAS were assessed by using the intraclass correlation coefficient (ICC) and strength of agreement for ordinal data by using \( \kappa \) statistics. The intrarater ICC for the mean GRBAS was 0.914 (95% confidence interval [CI] 0.698–0.977), G score agreement was 85.0% and \( \kappa \) 0.571. The intrarater ICC for the mean GRBAS was 0.846 (95% CI 0.457–0.958), G score agreement was 84.9% and \( \kappa \) 0.567. These results demonstrated moderate inter- and intrarater agreement, thus reliability of the GRBAS measures were considered acceptable for this study.

Cases were grouped according to their G score as normal to mild (0 to 1) or moderate to severe (2 to 3) dysphonia. Group comparisons were made by using Mann-Whitney tests for continuous measures and \( \chi^2 \) tests for categorical comparisons. Multivariable logistic regression was conducted to assess the influence of demographic and intubation characteristics on the severity of dysphonia. Significant factors were summarized by using odds ratios (ORs) and 95% CIs. SPSS 18.0 statistical software (IBM SPSS Statistics, IBM Corporation, Chicago, IL) and LogXact (Cytel Inc., Cambridge, MA, 2007) were used for data analysis. All hypothesis tests were 2-sided and \( P < .05 \) were considered statistically significant.

**RESULTS**

There were 317 infants live born in Western Australia between 22\(^{0}\) and 24\(^{6}\) weeks’ gestation from 1996 to 2004, of whom 251 (79%) were admitted to the tertiary NICU. Of these, 241 were inborn at the tertiary center and 10 were born elsewhere, requiring neonatal emergency transport services to the tertiary center. There were 154 survivors (61%) to the present date (22 weeks 7/24 [29%], 23 weeks 43/88 [49%], 24 weeks 104/139 [75%]).

Twenty-four children were excluded based on distant residence (\( n = 13 \)) or disability likely to preclude successful assessment (\( n = 11 \)), leaving a target population of 130. Nine were untraceable. Seventy-nine agreed to attend, of whom 71 attended, and 67 completed valid assessments. Baseline and intubation characteristics comparing those assessed with those who were not, are shown in Table 1.

All infants were intubated at birth, and subsequently for a median total duration of 45 days. Those tested were ventilated for 11 fewer days than the group not tested (\( P = .013 \)). The shortest period of intubation was for 1 day and the longest was for 113 days. Maximum ETT diameters for most were either 2.5 or 3.0 mm, whereas the maximum ETT ratio varied from 3.0 to 7.2, the latter for a 415-g infant reintubated at 15 days of age with a 3.0-mm ETT. Although the median number of intubations was 5 per infant, 1 child had 16 intubations. Only 4 (6%) of the 67 preterm tested children demonstrated a normal G score, and a further 24 (36%) had mild hoarseness, as did 41% of the term group (Table 2); however, 36 (54%) showed moderate and 3 (5%) severe hoarseness, compared with just 1, in the moderate range, in the term group. There was a highly significant difference in the mean of the total GRBAS values between the groups.

The parent-reported scores for the preterm group on the pVHI were far higher than the expected scores for both normophonic children (≤2) and for our term group (\( P < .001 \)). There was also a clear association of higher scores in children with moderate or
severe voice abnormality on the GRBAS measure (Fig 1). Parents of children with moderate to severe voice abnormality reported higher impact on quality of life in each of the 3 areas (all \( P \) values < .001).

On univariate analyses, moderate-severe voice abnormality was associated with female gender, longer duration of intubation, higher frequency of intubation, and maximum tube size (Table 3). There were no differences in the frequency or duration of intubation, maximum tube size, or maximum ETT ratio between boys and girls (all \( P \) values > .05).

Simultaneous modeling of demographic and intubation characteristics showed an increased odds of moderate-severe voice abnormality for children who had a history of more than 5 intubations (OR 6.96, 95% CI 2.07–23.40, \( P = .002 \)) and for girls relative to boys (OR 3.46, 95% CI 1.12–10.62, \( P = .030 \)). Maximum tube size and total duration of intubation were not significant in the multivariable model (\( P \) values .438 and .810 respectively).

**DISCUSSION**

Although voice disorder of a mild degree is common in healthy children,\(^{24}\) the rate of moderate to severe voice abnormality, particularly hoarseness, in a school age population of extremely preterm neonates has been shown in this study to be unexpectedly high, at 58%. Factors significantly associated with the presence of a voice disorder included a history of more than 5 episodes of endotracheal intubation and female gender. Garten et al\(^{19} \) found dysphonia at 12 months of age in 6.6% in ELBW infants; however, their use of an “in-house” perceptual dysphonia score rather than a standardized auditory perceptual measure, along with a much younger study population, makes it difficult to draw comparisons between their results and those obtained by this study.

Unilateral vocal fold paralysis has been found in more than half of ELBW infants undergoing surgical ligation of the PDA,\(^{13} \) and this has been shown to be a persisting problem into adulthood\(^{25} \) with voice problems and hoarseness as the most frequent findings. However, surgical ligation of the PDA is rarely carried out in our service, and occurred in only 4 of the 154 survivors in this study, and therefore did not play a significant part in the voice problems of this cohort.

The results of this study suggest that laryngeal injury as a result of endotracheal intubation in the neonatal period may persist into childhood, in contrast to previous reports that mild-moderate injuries sustained from intubation may self-resolve.\(^{12} \) Histologic changes affecting the cricoarytenoid joint have been shown after intubation\(^{17} \) and other studies have indicated the incidence of laryngeal injury as high as 95%.\(^{8,9,12} \) Supporting our findings that children born extremely preterm and intubated frequently are at increased risk for voice disorders during later childhood. Several studies of laryngeal injury after neonatal intubation have found an
association between prolonged duration of intubation and the presence of injury. Fan et al found that intubation for longer than 7 days was associated with moderate-major injury to the larynx visualized on laryngoscopy at least 1 week post extubation. However, Albert et al performed a similar prospective study and found no significant association between any of the intubation factors postulated by other authors as contributing to laryngeal injury. Our study suggests that a history of >5 episodes of endotracheal intubation was more predictive of a voice disorder in childhood than the duration of intubation. The number of intubation episodes has been found to be associated with laryngeal injury by several authors. Each episode can be associated with trauma to the airway; however, it is often difficult to examine the contribution of traumatic intubation because of inconsistencies of documentation. Complicated intubation, defined as >2 attempts, bleeding postprocedure, or visualization of swollen anatomic structures, was the only clear factor found in the study by to be a significant predictor of vocal dysfunction in extremely preterm neonates assessed at 1 year of age. In this study, we were not able to reliably determine from chart review the degree of difficulty of intubation. In our study, assessing the size of the ETT relative to the size of the infant, by using the maximum ETT ratio, did not show a correlation with the subsequent presence of a voice disorder, despite previous hypotheses proposed to that effect. A follow-up study performed by Sherman and Nelson that examined the rates of subglottic stenosis after the implementation of an ETT protocol dictating the “appropriate size” of ETT, reported the incidence of subsequently diagnosed stenosis to be significantly reduced. Appropriate size was determined as being the tube diameter divided by the infant’s gestational age in weeks being less than 0.1. The size of ETT chosen by a treating neonatologist is influenced by a number of factors, however, primarily to provide effective ventilation.

The presence of voice disorders in childhood is recognized as having an impact on long-term voice use, influencing social and academic function in

![FIGURE 1](https://example.com/figure1.png)

**FIGURE 1**

pVHI scores for functional, physical, and emotional subscales in the normal to mild (n = 28) and moderate to severe (n = 39) voice disorder groups.

**TABLE 3 Voice Abnormality: Univariate Associations With Birth and Intubation Characteristics**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Grade of Hoarseness</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal-Mild n = 28</td>
<td>Moderate-Severe n = 39</td>
</tr>
<tr>
<td>Gestational age at birth, wk</td>
<td>(24–24; 22–24)</td>
<td>(23–24; 22–24)</td>
</tr>
<tr>
<td>Birth weight, g</td>
<td>(589–728; 470–885)</td>
<td>(574–686; 445–790)</td>
</tr>
<tr>
<td>Male gender</td>
<td>(64)</td>
<td>(36)</td>
</tr>
<tr>
<td>Total intubations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1–5</td>
<td>(82)</td>
<td>(41)</td>
</tr>
<tr>
<td>&gt;5</td>
<td>(18)</td>
<td>(58)</td>
</tr>
<tr>
<td>Maximum tube size, mm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.5</td>
<td>(64)</td>
<td>(59)</td>
</tr>
<tr>
<td>3.0/3.5</td>
<td>(36)</td>
<td>(61)</td>
</tr>
<tr>
<td>Maximum ETT ratio</td>
<td>(3.6–4.5; 3.2–5.3)</td>
<td>(4.0–4.5; 3.6–7.2)</td>
</tr>
<tr>
<td>Total ETT duration, d</td>
<td>(17–42; 2–70)</td>
<td>(33–61; 17–92)</td>
</tr>
</tbody>
</table>

Data represent median (interquartile range; range) or n (%), as appropriate.
childhood and influencing career options in adulthood. Many of the extremely preterm children in this study have ongoing difficulties with health, development, and learning and the pVHI results indicate that their voice abnormalities result in additional functional and emotional difficulties.

We demonstrated an increased risk of voice disorder in female children, but were not able to identify any difference in neonatal or intubation risk factors between boys and girls and we do not have an explanation for this finding at present. There were several limitations of this study. First, the population of infants delivered at <25 weeks is a very small proportion of intubated newborns and further studies are needed to determine the extent of voice disorders in children of older gestational ages who require intubation. Second, despite the good inter- and intrarater reliability of the GRBAS tool in this and other studies, this was a subjective assessment in which the examiners were not blinded to the identity or the gestational age at birth of the participants; however, neonatal data were not available to the examiners at the time of voice assessment. Third, the group studied were <50% of the identified target population and therefore may not be representative of all survivors of preterm birth at <25 weeks from 1996 to 2004; however, we are able to identify all survivors at these gestations in Western Australia, and have not demonstrated major differences in neonatal morbidity, intubation history, or ventilation between participating and nonparticipating preterm children. The longer duration of ventilation in the untested children may reflect our exclusion criteria of more disabled children who may have been sicker in the neonatal period. Selection bias could lead this study to either overestimate or underestimate the true prevalence of voice disorder in this population; however, even in the unlikely event of all untested children born at <25 weeks having a normal voice, this study would indicate a rate of voice disorder far higher than has previously been recognized. Finally, this study did not correlate voice abnormality with an assessment of laryngeal pathology. Some of our more severely affected children have had detailed laryngeal assessments, but systematic assessment of moderate/severe voice abnormality including videostroboscopy may be useful for future studies.

REFERENCES


Voice Abnormalities at School Age in Children Born Extremely Preterm
Noel French, Rona Kelly, Shyan Vijayasekaran, Victoria Reynolds, Jodi Lipscombe, Ali Buckland, Jean Bailey, Elizabeth Nathan and Suzanne Meldrum

*Pediatrics* 2013;131;e733; originally published online February 18, 2013;
DOI: 10.1542/peds.2012-0817

| Updated Information & Services | including high resolution figures, can be found at: /content/131/3/e733.full.html |
| References | This article cites 29 articles, 3 of which can be accessed free at: /content/131/3/e733.full.html#ref-list-1 |
| Citations | This article has been cited by 3 HighWire-hosted articles: /content/131/3/e733.full.html#related-urls |
| Subspecialty Collections | This article, along with others on similar topics, appears in the following collection(s): Fetus/Newborn Infant /cgi/collection/fetus:newborn_infant_sub |
| Permissions & Licensing | Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at: /site/misc/Permissions.xhtml |
| Reprints | Information about ordering reprints can be found online: /site/misc/reprints.xhtml |