Performance of commonly used respiratory questionnaire items in a cohort of infants born preterm

Elizabeth Boggs¹, Nori Minich², Anna Maria Hibbs²,³

¹School of Medicine, Case Western Reserve University, Cleveland, USA
²Department of Pediatrics, Case Western Reserve University, Cleveland, USA
³Division of Neonatology, Rainbow Babies and Children’s Hospital, Cleveland, USA

Email: Elizabeth.Boggs@case.edu

Received 10 July 2013; revised 9 August 2013; accepted 17 August 2013

ABSTRACT

Background: Items from respiratory questionnaires validated in older children are often used in research studies of preterm infants, although they have not been validated in this population. We aimed to assess both test-retest reliability and convergent validity of a group of commonly used respiratory questionnaire items in a cohort of preterm infants. Methods: The health status of 300 preterm infants was assessed by telephone questionnaire as part of a prospective cohort study. The questionnaire items analyzed in this study included six commonly used respiratory questions. The questionnaire responses used in this analysis were from the telephone follow-up in this cohort at six months of age adjusted for prematurity. A repeat interview one to two weeks after this interview was performed in a subset of subjects to assess test-retest reliability. The convergent validity of the respiratory items was also assessed by calculating the associations among the responses to the respiratory questions. Results: A total of 43 infants were singletons that met the criteria for test-retest reliability analysis. All of the respiratory questions demonstrated fair to strong test-retest reliability. Among 206 respondents, respiratory questionnaire items also demonstrated strong convergent validity, in that caretakers reporting wheezing or whistling in the chest were significantly more likely to also report other respiratory events. Conclusions: This selection of standard respiratory questionnaire items performed well for research purposes in this population.

Keywords: Respiratory Questionnaire; Preterm; Wheeze

1. INTRODUCTION

Persistent wheezing in the first year of life is a common and long-lasting complication of prematurity [1,2,3-7]. Preterm infants born at less than 37 weeks gestational age (GA) are more likely than term-born infants to experience wheezing in infancy and also to develop asthma [2,4,8-12]. While several questionnaires assessing respiratory health have been well-validated in older children, their performance in infants, and in particular preterm infants, have not been well assessed [13-19].

Non-invasive questionnaire-based assessments of the respiratory outcomes of neonatal intensive care unit (NICU) graduates often borrow from validated questionnaires designed for older children, often those with asthma. For instance, in the follow-up of trials testing interventions to decrease bronchopulmonary dysplasia, parent-perceived wheezing is often assessed by asking about “wheezing and whistling in the chest”, a phrasing borrowed from questionnaires such as the International Study of Asthma and Allergies (ISAAC) questionnaire [17,18] and American Thoracic Society child questionnaire ATS-DLD-78-C [19]. In addition, parental reports of medical interventions, such as hospitalization, physician diagnoses, or prescription of medications, are also frequently used as markers of the long-term pulmonary status of infants born preterm [20-26]. Such questions have been used in the follow-up assessments of major multi-center trials studying preterm infants (e.g. NO CLD [NCT00006401] [14,20] and SUPPORT [NCT-00233324]) [21,22,27]. However, the reliability and validity of these questions in the premature population have not been established.

The objective of this study was to assess the performance of a group of commonly used respiratory questionnaire items in a preterm cohort. We aimed to both assess the test-retest reliability and explore the convergent validity of this set of questionnaire items. We hypothesized that the items would demonstrate good test-retest reliability and that the responses to questions about respira-
tory morbidity, such as parental reports of “wheezing or whistling in the chest” or inhaled medication use, would be highly correlated with each other.

2. METHODS

2.1. Population

As part of a prospective cohort study of 300 preterm infants, “Gastrointestinal Risk Factors for Wheezing in Premature Infants (GRWPI)”, health status was assessed by telephone questionnaire at 3, 6, 9, and 12 months of age adjusted for prematurity. This study enrolled infants born at 280/7 - 340/7 GA. Infants with bronchopulmonary dysplasia (supplemental oxygen requirement >28 days), major anomalies, severe neurological injury, or a history of necrotizing colitis were excluded from the study. All participants were born at or transferred to Rainbow Babies and Children’s Hospital.

2.2. Performance of Questionnaire Items

This analysis of the performance of the respiratory items on the questionnaire focused on the assessment at 6 months adjusted age. At this time point, a repeat interview was planned one to two weeks after the primary interview in the first 60 infants that were followed. The analysis of test-retest reliability was restricted to singletons for whom the same caretaker participated in both interviews.

The questionnaire items analyzed in this study included the following:

- “Has your child had wheezing or whistling in the chest?”
- “Has a doctor diagnosed your child with wheezing or asthma?”
- “Has your child had a respiratory infection, such as a cold, bronchiolitis, or pneumonia?”
- “Has your child been diagnosed with RSV, or Respiratory Syncytial Virus?”
- “Has your child been seen in the emergency room for a breathing problem?”
- “Has your child been admitted to the hospital for a breathing problem?”

Answer choices were “yes,” “no,” and “unsure”. “Unsure” answers were treated as missing.

2.3. Convergent Validity

Odds of reporting a concurrent respiratory event given positive responses to another respiratory questionnaire item were calculated to assess internal consistency and convergent validity. An additional non-respiratory questionnaire item “In the past week, has your child refused a feed when hungry?” was included as a control.

2.4. Statistical Analysis

Test-retest reliability was assessed using both Cohen’s kappa and percent agreement. Associations between markers of respiratory morbidity (e.g. hospitalization and medication use) were assessed with odds ratios, Chi square’s, and Fischer’s exact tests. Statistical analyses were performed with Stata version 9.2 (StataCorp, College Station, TX).

3. RESULTS

Two hundred seventy nine infants were followed at the six month time point (93% follow up). Of these, 206 were singletons meeting criteria for the analysis. The chronological age of participants at this interview ranged from 206 to 278 days (Table 1).

Test and retest respiratory interviews were conducted...
in 61 infants, as opposed to the goal of 60, due to the presence of twins in the sample. Among the 61 infants, 43 met criteria for the test-retest reliability analysis. Demographic characteristics for the test-retest subset were generally representative of the whole cohort (Table 1). Repeat administrations of the questionnaire were done 6 - 19 days after the first administration.

In this relatively healthy population of preterm infants, several of the questions had a low rate of positive responses. Doctor diagnoses of wheezing or asthma, RSV infections, ER visits for breathing problems, and hospital admissions for breathing problems were the least frequently reported markers of respiratory morbidity.

Wheezing and whistling in the chest and respiratory infections had the highest reported rates (Table 2).

All of the respiratory questions demonstrated fair to strong test-retest reliability (Table 2). The strongest agreement existed for doctor diagnosis of wheezing or asthma, hospital admissions for breathing problems, and RSV infections. ER visits for a breathing problem and wheezing and whistling in the chest also demonstrated substantial agreement.

Convergent validity was assessed by calculating the odds of a caretaker answering an item affirmatively if they had also answered another item affirmatively (Table 3). Notably, caretakers reporting parental recognition

Table 2. Kappa scores and percent agreement for the test-retest reliability of the respiratory questionnaire.

<table>
<thead>
<tr>
<th>Respiratory Questions</th>
<th>Percent that Responded “Yes” in the First Interview</th>
<th>Kappa</th>
<th>Percent Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Has your child had a respiratory infection?</td>
<td>53.7%</td>
<td>0.40</td>
<td>69.8%</td>
</tr>
<tr>
<td>Has your child had wheezing or whistling in the chest since our last interview?</td>
<td>38.5%</td>
<td>0.73</td>
<td>88.4%</td>
</tr>
<tr>
<td>Has your child been seen in the ER for a breathing problem?</td>
<td>14.2%</td>
<td>0.88</td>
<td>97.8%</td>
</tr>
<tr>
<td>Has a doctor diagnosed your child with wheezing or asthma?</td>
<td>7.8%</td>
<td>1.0</td>
<td>100%</td>
</tr>
<tr>
<td>Has your child been admitted to the hospital for a breathing problem?</td>
<td>6.3%</td>
<td>1.0</td>
<td>100%</td>
</tr>
<tr>
<td>Has your child been diagnosed with RSV?</td>
<td>4.9%</td>
<td>1.0</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 3. Internal consistency and convergent validity. The odds ratios for reporting a respiratory event given a report of another event are listed.

<table>
<thead>
<tr>
<th>Odds of Reporting Respiratory Event OR [95% CI] p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wheezing or Whistling in the Chest</td>
</tr>
<tr>
<td>--------------------------------------------------------</td>
</tr>
<tr>
<td><strong>p &lt; 0.0001</strong></td>
</tr>
<tr>
<td>4.34</td>
</tr>
<tr>
<td>[2.34, 8.06]</td>
</tr>
<tr>
<td>p &lt; 0.0001</td>
</tr>
<tr>
<td>1.50</td>
</tr>
<tr>
<td>[0.523, 4.29]</td>
</tr>
<tr>
<td>p = 0.0001</td>
</tr>
<tr>
<td>3.92</td>
</tr>
<tr>
<td>[1.52, 10.1]</td>
</tr>
<tr>
<td>p &lt; 0.0028</td>
</tr>
<tr>
<td>18.1</td>
</tr>
<tr>
<td>[4.37, 75.3]</td>
</tr>
<tr>
<td>p &lt; 0.0001</td>
</tr>
<tr>
<td>6.30</td>
</tr>
<tr>
<td>[1.95, 20.4]</td>
</tr>
<tr>
<td>p &lt; 0.0006</td>
</tr>
<tr>
<td>23.3</td>
</tr>
<tr>
<td>[5.99, 90.8]</td>
</tr>
<tr>
<td>p &lt; 0.0001</td>
</tr>
<tr>
<td>52.3</td>
</tr>
<tr>
<td>[11.0, 248]</td>
</tr>
<tr>
<td>p &lt; 0.0001</td>
</tr>
<tr>
<td>1.39</td>
</tr>
<tr>
<td>[0.369, 5.22]</td>
</tr>
<tr>
<td>p = 0.7995</td>
</tr>
</tbody>
</table>

*p Odds ratio could not be calculated because there were no infants for whom the parents reported a diagnosis of wheezing but no wheezing symptoms; **Odds ratio could not be calculated because there were no infants for whom the parents reported a diagnosis of RSV but no wheezing symptoms; ***Odds ratio could not be calculated because there were no infants for whom the parents reported a respiratory hospitalization but no respiratory infection.
of wheezing or whistling in the chest were significantly more likely to also report a doctor diagnosis of wheezing, respiratory infection, RSV infection, ER visit for a breathing problem, hospitalization, and use of inhaled or use of oral steroids to address a breathing problem. Refusal to feed was significantly correlated with wheezing and whistling in the chest and doctor diagnosis of RSV infection, though overall it was not strongly correlated with this set of respiratory questions.

4. CONCLUSIONS

We assessed the performance of commonly used respiratory questionnaire items in a cohort of relatively healthy preterm infants. All of the respiratory questions demonstrated fair to strong test-retest reliability. Furthermore, these items demonstrated convergent validity; the markers of respiratory severity assessed were strongly associated with each other. Taken together, these results suggest that these standard respiratory questions perform well for research purposes in this population.

In the analysis of test-retest reliability, respiratory events with lower rates of positive responses tended to demonstrate better agreement. This is likely due to the fact that the less frequent events, such as hospitalization, may also be more salient for the parent. However, rare outcomes are non-ideal primary endpoints for research studies due to their impact on sample-size. Furthermore, hospitalization may be driven by a number of non-medical factors [28-35]. The more frequently reported parental recognition of symptomatic wheezing or whistling in the chest may be an outcome prone to more “noise” from misclassification; however, in real-world clinical interactions, physicians often rely on parental reports of symptom burden to drive care. Thus, studies wishing to assess respiratory morbidity by parental questionnaire benefit by capturing several complementary outcomes, which, as we have shown, do demonstrate strong convergent validity.

Interestingly, doctor diagnosis of wheezing or asthma performed very well in terms of test-retest reliability, yet it was the only item that was not significantly associated with other respiratory events in the exploration of convergent validity. One possible reason for this outcome is that the question “Has a doctor diagnosed your child with wheezing or asthma?” is a compound question, in that wheezing and asthma are two different diagnoses. A compound question was asked instead of a traditional wording validated in older children about doctor-diagnosed asthma due to reluctance of many practitioners to differentiate between wheezing and asthma in infants less than one year old. However, this may have weakened the performance of the question. Another possible reason for this outcome is that it reflects a true lack of association between doctor diagnosis and the other items assessed, perhaps due to a different threshold for using or recalling symptomatic treatment versus the application of a diagnostic label. Perhaps unsurprisingly, refused feed was correlated with wheezing and whistling in the chest and diagnosis of RSV infection, both indicators of symptomatic illness. Overall, this question did not correlate significantly with most of the respiratory questions, an indicator that correlation among respiratory items was not primarily driven by a general tendency of parents to affirm symptoms.

This study has several limitations. While these respiratory questions perform well in terms of test-retest reliability and convergent validity, it is noteworthy that the responses are not matched against medical records, thus we cannot assess the accuracy of these results. Furthermore, this analysis is a limited assessment of the performance of a series of commonly used respiratory questionnaire items. While it represents a first step in the assessment of the performance of these items in the preterm population, it is not a full validation of a free-standing questionnaire.

Nevertheless, this study offers reassurance that questionnaire items commonly used to assess respiratory outcomes in preterm infants perform well in terms of test-retest reliability and convergent validity. Reliable and valid non-invasive measures are ideal assessment tools in a medically fragile population in whom more invasive testing of pulmonary status may be both medically risky and expensive. Furthermore, respiratory symptoms and resource utilization may reflect the experience of patients and families better than physiologic measures. Though this study is an initial step in fully validating a questionnaire in this population, so far the performance of these commonly used items suggests that their use in studies of preterm infants is sound. Future work is needed to develop a fully validated questionnaire to standardize the assessment of respiratory status in preterm infants.

5. ACKNOWLEDGEMENTS

Other Contributions: We acknowledge the funding received from the National Institutes of Health [K23 HD056299].

REFERENCES


or conventional ventilation in babies of gestational age 28 weeks or less: Respiratory and neurological outcomes at 2 years. *Archives of Disease in Childhood Fetal and Neonatal Edition*, 91, F320-F326. doi:10.1136/adc.2005.079632


