Reliability of a Noninvasive Measure of $\dot{V}/\dot{Q}$ Mismatch for Bronchopulmonary Dysplasia

Nicolas Bamat¹, Sarvin Ghavam¹, Yumei Liu², Sara B. DeMauro¹, Erik A. Jensen¹, Robin Roberts³, Bradley A. Yoder⁴, and Haresh Kirpalani¹

¹Division of Neonatology, The Children's Hospital of Philadelphia, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania; ²Department of Neonatology, Guangdong Academy of Medical Science, Guangdong General Hospital, Guangzhou, China; ³Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada; and ⁴Department of Pediatrics, University of Utah School of Medicine, Salt Lake City, Utah

Abstract

Rationale: Currently used definitions of bronchopulmonary dysplasia (BPD) lack a continuous measure of disease severity.

Objectives: To determine if an indirect measure of $\dot{V}/\dot{Q}$ mismatch is reliable when simplified to facilitate more widespread use for grading disease severity in BPD at 36 weeks postmenstrual age.

Methods: We used prospectively collected data from 32 preterm infants undergoing an oxygen reduction test at 36 weeks postmenstrual age to perform a simplified indirect assessment of $\dot{V}/\dot{Q}$ mismatch for each infant. Independent raters applied the model, and interrater reliability for a quantitative measure of mismatch was measured by intraclass correlation coefficient. A receiver operating characteristic curve evaluated the impact of increasing degrees of $\dot{V}/\dot{Q}$ mismatch on diagnosing BPD as defined by oxygen reduction test failure.

Measurements and Main Results: Concordance for the quantitative measure of $\dot{V}/\dot{Q}$ mismatch between independent raters improved from 0.72 (confidence interval [CI], 0.48–0.86) to 0.93 (CI, 0.87–0.96) after refinement of instructions for applying the simplified model. Higher degrees of mismatch were increasingly predictive of oxygen reduction test failure, with a receiver operating characteristic curve analysis area under the curve of 0.83 (CI, 0.68–0.99; $P = 0.03$).

Conclusions: A simplified indirect measure of $\dot{V}/\dot{Q}$ mismatch for diagnosing and grading disease severity in BPD has high reliability and can be performed with data obtained during a standard oxygen reduction test. This should facilitate more widespread investigation of this model as a technique for characterizing BPD severity.

Keywords: premature infant; chronic lung disease; definition; interrater reliability; disease severity

(Received in original form October 10, 2014; accepted in final form February 23, 2015)

Supported by National Institutes of Health grant U-0HD053124-01 (B.A.Y.).

The views expressed in this article do not communicate an official position of the National Institutes of Health.

Author Contributions: All authors reviewed and approved the submission of this version of the manuscript and take full responsibility for the manuscript. N.B. contributed to the conception and design of the work; acquisition, analysis, and interpretation of the data; and was responsible for the primary drafting and revising of the manuscript. S.G. contributed to the design of the work, acquisition of the data, and critically revised the work. Y.L. contributed to the acquisition of the data and critically revised the work. S.B.D. contributed to the analysis and interpretation of data for the work and critically revised the work. E.A.J. contributed to the acquisition of the data, the analysis and interpretation of the data for the work, and critically revised the work. R.R. contributed to the analysis and interpretation of the data for the work and critically revised the work. B.A.Y. contributed to the acquisition of the data and critically revised the work. H.K. contributed to the conception and design of the work, analysis and interpretation of the data, and critically revised the work. All contributions were substantial.

Correspondence and requests for reprints should be addressed to Haresh Kirpalani, B.M., M.Sc., Division of Neonatology, Children's Hospital of Philadelphia, 34th and Civic Center Boulevard, 2nd Floor, Main, Philadelphia, PA 19104. E-mail: kirpalanih@email.chop.edu

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org

Ann Am Thorac Soc Vol 12, No 5, pp 727–733, May 2015

Copyright © 2015 by the American Thoracic Society

DOI: 10.1513/AnnalsATS.201410-462OC

Internet address: www.atsjournals.org

Bronchopulmonary dysplasia (BPD) is a common complication of prematurity and a strong predictor of poor neurodevelopmental outcomes (1). Consequently, it is a frequently used

outcome measure in neonatal clinical trials (2). Despite this, limitations in the definitions for BPD remain (3).

BPD is most commonly defined as a need for supplemental oxygen at 36 weeks

postmenstrual age (PMA) (4). However, variable provider use of supplemental oxygen limits the objectivity of this definition and contributes to the wide variability in the incidence of BPD observed

between centers (5). To address this, the oxygen reduction test (ORT) (6) applies a more objective definition of demonstrated oxygen need to diagnose the presence of disease. This is now commonly used as the defining criterion in trials where BPD is an outcome measure (2). These applications of the ORT describe a simple dichotomous presence or absence of BPD, without reporting a measure of severity. However, because this disease has a wide spectrum of respiratory impairment (7), this limits the usefulness of this definition.

An additional physiologic definition of BPD proposes a noninvasive quantitative measure of the degree of $\dot{V}/\dot{Q}$ mismatch. This approach uses sequential measurements of arterial oxygen saturation (SpO_2) as a function of delivered inspired oxygen pressure (PI_{O_2}) to generate infant-specific curves for comparison against normative values (8). Quantifiable characteristics of the curves provide a measure of disease severity for each infant.

In contrast to the ORT, the $\dot{V}/\dot{Q}$ model is not widely used. This may be due to the relatively cumbersome methodology of the model as originally described (8, 9). We simplified the methodology of the model to allow its use with data available from a standard ORT procedure. Our primary objective was to determine if this simplified model was reliable, as defined by a high

degree of interrater reliability. A secondary objective was to assess whether progressively higher degrees of disease severity, as measured by the simplified model, could increasingly predict failure on the ORT. We hypothesized that if the objectives were met, the simplified model could potentially receive more widespread use to measure disease severity in BPD. Some results of these studies have been previously reported in the form of an abstract (10).

Methods

Population

Prospectively collected data were used from 32 very low birth weight (<1,500 g) preterm infants admitted to the neonatal intensive care units of the University of Utah University Hospital (Salt Lake City, UT) and Intermountain Medical Center (Murray, UT). An ORT was performed when the following inclusion criteria were met: a PMA of 36 to 37 weeks and maintenance of SpO_2 greater than 90% in the preceding 24 hours while receiving supplemental oxygen less than an effective FI_{O_2} of 0.36 (equivalent to 0.30 after correction for the barometric pressure of 644 mm Hg at our study sites; Salt Lake City, UT: 1,311 m, Murray UT: 1,320 m). Temporary increases in FI_{O_2} for transient decreases in SpO_2 associated with apnea,

bradycardia, feeding, or manipulations were allowed as long as the infant returned to baseline within 1 hour. Exclusion criteria were a need for supplemental oxygen greater than effective FI_{O_2} greater than 0.36 or positive pressure support. Institutional review boards approved all interventions. Parental consent was obtained for the original intervention and data collection; it was waived for this secondary, deidentified analysis.

The ORT

To conduct the standard ORT, we followed procedures described by Walsh and colleagues (6). Infants receiving supplemental oxygen with an effective fraction of inspired oxygen (FI_{O_2}) between 0.21 and 0.30 underwent a timed, step-wise reduction of oxygen to room air, with continuous SpO_2 monitoring. The effective FI_{O_2} was determined based on the liter per minute nasal cannula flow, infant weight in kilograms, and the delivered FI_{O_2} , as described by Benaron and Benitz (11). Starting with the infant's baseline support, step-wise reductions in flow and FI_{O_2} were performed every 5 minutes until either failure of the ORT or progression to room air without failure. SpO_2 values were recorded at 1-minute intervals. Failure of the ORT was defined as an SpO_2 less than 90% for 5 consecutive minutes or less than 80% for 15 seconds with a good signal waveform. Infants who failed the ORT were

Figure 1. Shunt versus $\dot{V}/\dot{Q}$ mismatch. Plots of oxyhemoglobin saturation (SpO_2 , %) against inspired oxygen pressure (PI_{O_2} , kPa). (A) Increasing shunt from 0 to 40% lowers the position of the upper part of the curve. (B) Reducing $\dot{V}/\dot{Q}$ from 0.8 to 0.1 shifts the curve to the right. The right shift of each PI_{O_2} – SpO_2 curve from position of the dissociation curve (dashed line) is the $PI_{O_2} - Pa_{O_2}$ difference (kPa), which includes Pa_{CO_2} /respiratory gas exchange rate. The 0.8 curve represents the normal adult curve, which intercepts a PI_{O_2} of 21 kPa (vertical line) at 97% SpO_2 . Reprinted by permission from Reference 8.

classified as having BPD. For infants successfully reaching room air, a continuous 15-minute period with Sp_{O_2} greater than 96% or 30 minutes without failure was needed to pass the ORT.

Simplification of the $\dot{V}/\dot{Q}$ Model

The indirect $\dot{V}/\dot{Q}$ model has been described in detail previously (8). The model characterizes the relationship between Pi_{O_2} and Sp_{O_2} in infants on supplemental

oxygen. Delivered Pi_{O_2} is altered in a step-wise fashion and resulting changes in Sp_{O_2} recorded. Multiple paired values are plotted graphically, with Sp_{O_2} on the ordinate (y) and Pi_{O_2} in kPa on the abscissa (x). Plotted points are connected by a best-fit curve. Two primary impairments of gas exchange have predictable effects on the shape of the curve (Figure 1). The family of curves in Figure 1A demonstrates the flattening of

the curve seen with an increasing right-to-left shunt. Figure 1B demonstrates the progressive right shift of the curve observed as the $\dot{V}/\dot{Q}$ ratio decreases. The degree of right shift uses the oxygen dissociation curve in an “ideal lung” as a reference and reflects the inspired oxygen pressure to arterial partial pressure of oxygen ($Pi_{O_2} - Pa_{O_2}$) difference. This is determined by the reduction in $\dot{V}/\dot{Q}$ ratio and by the alveolar

Figure 2. $\dot{V}/\dot{Q}$ model application. Representative screen shot image of software available to facilitate use of the $\dot{V}/\dot{Q}$ model. Numerical steps provide an overview of the process for obtaining degree of right shift in kPa. Each reduction step generates a paired value of inspired oxygen pressure (Pi_{O_2}) and arterial oxygen saturation percentage (Sp_{O_2} , %), represented by a single point. Multiple plotted points are connected by a selected best-fit curve, achieved by moving the overlying curves to best approximate all plotted points. This imprecise selection process allows inherent subjectivity and the potential for disagreement between independent raters. Degree of right shift in kPa is calculated by the software after selection of a best-fit curve, uses the oxygen dissociation curve in an ideal lung (leftmost curve within graph) as a reference standard, and reflects the inspired oxygen pressure to arterial partial pressure of oxygen ($Pi_{O_2} - Pa_{O_2}$) difference for that infant. This reflects degree of $\dot{V}/\dot{Q}$ reduction and is proposed as a measure of disease severity in bronchopulmonary dysplasia. Adapted by permission from Reference 20.

partial pressure of carbon dioxide. Preterm infants needing supplemental oxygen at 36 weeks PMA predominantly show a right shift of the curve. Accordingly, this measure was proposed as a potential method for defining BPD, with the degree of right shift representing disease severity (8).

The original method to evaluate $\dot{V}/\dot{Q}$ mismatch used precise measurements of P_{iO_2} and Sp_{O_2} (8, 9). This required infants on nasal cannula to be instead placed in an oxyhood with oxygen measured via an oxygen analyzer. With each change in P_{iO_2} , the infant was stabilized for approximately 5 minutes, after which a single, stable Sp_{O_2} value was recorded (8). In this current study, we simplified the $\dot{V}/\dot{Q}$ model such that all necessary data could be obtained during an ORT. Rather than directly measure P_{iO_2} , effective F_{iO_2} was calculated at each successive reduction while infants remained on nasal cannula oxygen (11). At sea level (101.3 kPa), percent effective F_{iO_2} approximates P_{iO_2} in kPa, allowing substitution of equivalent absolute values to arrive at P_{iO_2} . As this study was conducted above sea level at a barometric pressure of approximately 644 mm Hg, P_{iO_2} values were divided by a correction factor of 1.2 (760/644 mm Hg) to arrive at the effective P_{iO_2} . Next, we calculated the mean Sp_{O_2} from values recorded every minute over the length of each oxygen reduction step, obviating the need for a 5-minute period of stabilization before measuring a single value. These simplifications enabled both the $\dot{V}/\dot{Q}$ model and the ORT to be applied using the same procedure.

For each infant, calculated effective P_{iO_2} was plotted against the mean Sp_{O_2} for each reduction step of the ORT, creating a series of plotted points, one for each reduction step. We then approximated a best-fit curve to the series of plotted points for each infant, using software (<http://www.noranaes.org/shuntcurves/shuntcurves.html>) created to facilitate the use of the $\dot{V}/\dot{Q}$ model (9). As the relationship between P_{iO_2} and Sp_{O_2} is influenced by hemoglobin concentration, the hemoglobin level measured closest in time to the ORT for each infant was documented and applied to the $\dot{V}/\dot{Q}$ model as per the provided software (Figure 2). Raters followed procedural instructions to minimize interpretational laxity (12) in selecting a best-fit curve and resulting degree of right shift in kPa. To achieve this, we used an iterative process to

identify instances of disagreement between two raters and refine the instructions. Three iterations of instructions were tested, leading to three successive evaluations of interrater reliability. The third and final evaluation was performed by one of the original raters and two additional naive raters. The final set of instructions is provided as an online data supplement. After demonstration of high concordance between raters, right shift values determined by one of the original raters were used to compare the $\dot{V}/\dot{Q}$ model to results of the ORT.

Statistics

Baseline and demographic characteristics were summarized with standard descriptive statistics. An intraclass correlation coefficient using a two-way mixed effects model for single measures assessed interrater reliability for right shift values. The degree of right shift between patients who passed and failed the ORT were compared with a Mann-Whitney *U* test for nonparametric continuous variables (Stata 12; StataCorp, College Station, TX). The sensitivity and specificity of the simplified $\dot{V}/\dot{Q}$ model for diagnosing BPD using ORT failure as a reference standard was

evaluated using a receiver operating characteristics (ROC) curve.

Results

Thirty-two infants were included, with a mean (range) birth weight of 808 (590–1,210) grams and gestational age of 26.0 (23.2–28.2) weeks. Four infants (12.5%) passed and 28 (87.5%) failed the ORT. One infant failed the ORT at baseline before a reduction step, with only a single Sp_{O_2} and P_{iO_2} pair being available. As the $\dot{V}/\dot{Q}$ model requires at least two separate data points to approximate a best-fit curve, this infant was excluded from the simplified $\dot{V}/\dot{Q}$ model, leaving 31 infants included in the remaining analysis. The simplified $\dot{V}/\dot{Q}$ model was applied as described above to 29 of the 31 subjects during the initial iteration and to all 31 subjects during the second and third iterative instructions.

Interrater reliability for the initial assessment of the degree of right shift was 0.72 (95% confidence interval [CI], 0.48–0.86; $P < 0.001$). Refinement of instructions for selection of a best-fit curve resulted in

Figure 3. Interrater reliability for simplified $\dot{V}/\dot{Q}$ model. Measurement of degree of right shift in kPa as calculated by rater A against the degree of right shift (kPa) as calculated by rater B. The line depicts exact agreement ($y = x$). Each patient is represented by a data point. Overall concordance with intraclass correlation coefficient was 0.93 (95% confidence interval, 0.87–0.96; $P < 0.001$), reflective of third and final iteration of instructions, as detailed in the online data supplement.

progressive improvement in concordance. The intraclass correlation coefficients for the second and third iterations were 0.83 (95% CI, 0.67–0.91; $P < 0.001$) and 0.93 (95% CI, 0.87–0.96; $P < 0.001$), respectively. The degree of interrater reliability between two of the three raters for the final set of instructions is displayed graphically in Figures 3 and 4. Figure 3 displays the concordance as a plot of the degree of right shift (kPa) for each patient between two independent raters, and Figure 4 shows a histogram of the absolute difference in degree of right shift for each patient case (13).

The mean right shift value for the cohort as a whole as determined by rater A was 9.9 kPa (SD, 1.5). A histogram displaying the distribution of right shift values for the study cohort is displayed in Figure 5. Infants passing the ORT had a lower right shift value (mean = 8.1 kPa; SD, 1.25), implying less $\dot{V}/\dot{Q}$ mismatch and less disease severity than those failing the test (mean = 10.1 kPa; SD, 2.3; $z = 2.13$; $P = 0.03$).

The specificity of the degree of right shift to diagnose BPD as defined by ORT failure increased progressively with greater right shift values. ROC curve analysis

resulted in a statistically significant area under the curve of 0.83 (95% CI, 0.68–0.99; $P = 0.03$) (Figure 6).

Discussion

This study simplified an existing diagnostic model for BPD that provides a continuous measure of disease severity, allowing its application through methods commonly used in clinical research: the ORT. Application of our pragmatic modifications resulted in high interrater agreement, suggesting this is a feasible and reproducible diagnostic approach. Additionally, by showing that greater right shift values increasingly predict failure of the ORT, we demonstrate validity through comparison to a current diagnostic standard.

Approaches for better defining BPD are needed (3). A definition that provides a continuous measure of severity is of particular importance for this disease presenting along a wide spectrum of severity. Preterm infants requiring respiratory support have variable degrees of impairment in pulmonary function, and an ideal definition of BPD would reflect these gradations. The NIH consensus definition

acknowledges this through the categorical classification of the disease as mild, moderate, or severe (14). Any dichotomous definition of a disease with a broad spectrum of pathology is open to the arbitrary inclusion of false positives or the exclusion of false negatives. Healthy preterm infants have saturations around 97% in room air (15), well above the threshold of 90% often used in the ORT. Pulmonary function testing abnormalities are present at 9 years of age in 29% of very low birth weight infants who did not require any supplemental oxygen at 36 weeks PMA (16, 17). These infants would have been categorized as not having BPD without requiring further testing by the ORT. This exemplifies the shortcomings of dichotomously applied definitions of BPD.

The usefulness of a definition of BPD that provides a continuous measure of disease severity is significant for clinical research studies. Higher disease severity may be more predictive of respiratory and neurodevelopmental impairments later in life. Furthermore, the use of a continuous outcome measure may facilitate identification of efficacy in neonatal interventions.

In our current study, the application of the modified $\dot{V}/\dot{Q}$ model was limited to a cohort of infants undergoing an ORT. Our pragmatic approach took advantage of the existence of the necessary data among this cohort of similar patients with small baseline differences in their degree of respiratory support. However, the $\dot{V}/\dot{Q}$ model could be applied to preterm infants along a broader spectrum of pulmonary impairment. This includes infants on room air maintaining Sp_{O_2} values above currently used thresholds (by applying supplemental oxygen and noting the impact of variable Pi_{O_2} on Sp_{O_2}) as well as infants requiring supplemental oxygen in excess of an effective Pi_{O_2} of 0.30 and those requiring positive pressure.

We acknowledge the existence of previously reported methods for quantifying disease severity in infants with BPD (18) and that within our own cohort, noting the effective Pi_{O_2} during the last tolerated reduction step would have provided a measure of severity for those infants failing the ORT. However, the $\dot{V}/\dot{Q}$ model offers a continuous measure of disease severity with a physiologic underpinning. The degree of right shift

Figure 4. Absolute right shift difference between raters. Histogram of the absolute difference in degree of right shift in kPa as calculated by two raters independently applying the simplified $\dot{V}/\dot{Q}$ model. Concordance reflective of third and final iteration of instructions, as detailed in the online data supplement.

Figure 5. Cohort distribution of right shift values. Histogram of the distribution of degree of right shift in kPa for full study cohort by rater A.

reflects the $P_{iO_2} - Pa_{O_2}$ difference and degree of V/Q reduction. The authors of the original model demonstrated that it is this

impairment in gas exchange that predominantly characterizes infants with BPD, thus providing the impetus for its use

Figure 6. Impact of degree of right shift on oxygen reduction test failure. Receiver operator characteristic curve showing impact of degree of right shift (RS) in kPa on predicting bronchopulmonary dysplasia as defined by failure of the oxygen reduction test; area under the curve, 0.83; 95% confidence interval, 0.68–0.99; $P = 0.03$.

(8). This physiologic approach allows discrimination from other factors contributing to hypoxemia, namely shunt fraction.

We acknowledge several limitations. Our analysis is based on a small sample size of 32 infants. Furthermore, only four subjects “passed” the ORT and were classified as not having BPD. Despite this, the area under the ROC curve was greater than 0.8. Larger studies are necessary to better quantify the right shift levels that correspond to the absence or presence of BPD using current definitions such as the ORT.

Our participating centers were above sea level, resulting in lower barometric pressures and P_{iO_2} values. Although we have accounted for this, the effective P_{iO_2} for our subjects was often less than that of room air at sea level. This resulted in curves starting at a lower x-value and with less absolute right shift values than one would expect at sea level. This does not impact our measure of interrater reliability or our comparison of increasing right shift values to the ORT within this cohort. However, it would not be appropriate to compare the absolute values of right shift reported here to similar populations at sea level.

A detailed set of instructions (see the online data supplement) was used to obtain the high interrater reliability reported in this study. This facilitated the reliable selection of a best-fit curve between independent raters. There exists an inherent subjectivity to selecting a best-fit curve, a phenomenon that has been observed in other studies requiring judgment of curvilinear outputs (12). Although not a shortcoming of our methods, software calculating the degree of right shift and circumventing this subjective selection would be useful.

In summary, we describe simplifications to an existing model for measuring disease severity in BPD. These changes greatly facilitate its application. We demonstrate a high level of concordance between raters and validity for the model when compared with a current diagnostic standard. Further evaluation of this model for characterizing disease severity among infants with evolving and established BPD is warranted. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

References

- Schmidt B, Asztalos EV, Roberts RS, Robertson CM, Sauve RS, Whitfield MF. Impact of bronchopulmonary dysplasia, brain injury, and severe retinopathy on the outcome of extremely low-birth-weight infants at 18 months: results from the trial of indomethacin prophylaxis in preterms. *JAMA* 2003;289:1124–1129.
- Kirpalani H, Millar D, Lemyre B, Yoder BA, Chiu A, Roberts RS. A trial comparing noninvasive ventilation strategies in preterm infants. *N Engl J Med* 2013;369:611–620.
- Bancalari E, Claure N. Definitions and diagnostic criteria for bronchopulmonary dysplasia. *Semin Perinatol* 2006;30:164–170.
- Van Marter LJ. Epidemiology of bronchopulmonary dysplasia. *Semin Fetal Neonatal Med* 2009;14:358–366.
- Payne NR, LaCorte M, Karna P, Chen S, Finkelstein M, Goldsmith JP, Carpenter JH. Reduction of bronchopulmonary dysplasia after participation in the Breathsavers Group of the Vermont Oxford Network Neonatal Intensive Care Quality Improvement Collaborative. *Pediatrics* 2006;118:S73–S77.
- Walsh MC, Wilson-Costello D, Zadell A, Newman N, Fanaroff A. Safety, reliability, and validity of a physiologic definition of bronchopulmonary dysplasia. *J Perinatol* 2003;23:451–456.
- Jobe AH. Good news for lung repair in preterm infants. *Am J Respir Crit Care Med* 2013;187:1043–1044.
- Quine D, Wong CM, Boyle EM, Jones JG, Stenson BJ. Non-invasive measurement of reduced ventilation:perfusion ratio and shunt in infants with bronchopulmonary dysplasia: a physiological definition of the disease. *Arch Dis Child Fetal Neonatal Ed* 2006;91:F409–F414.
- Rowe L, Jones JG, Quine D, Bhushan SS, Stenson BJ. A simplified method for deriving shunt and reduced VA/Q in infants. *Arch Dis Child Fetal Neonatal Ed* 2010;95:F47–F52.
- Bamat N, Ghavam S, Liu Y, DeMauro SB, Yoder BA, Kirpalani H. Is the non-invasive assessment of V/Q mismatch reliable and how does it compare with oxygen reduction test to diagnose bronchopulmonary dysplasia [abstract]? Presented at the Pediatric Academic Societies Annual Meeting. April 28-May 1, 2012, Boston, MA.
- Benaron DA, Benitz WE. Maximizing the stability of oxygen delivered via nasal cannula. *Arch Pediatr Adolesc Med* 1994;148:294–300.
- Mehta S, Stewart TE, MacDonald R, Hallett D, Banayan D, Lapinsky S, Slutsky A. Temporal change, reproducibility, and interobserver variability in pressure-volume curves in adults with acute lung injury and acute respiratory distress syndrome. *Crit Care Med* 2003;31:2118–2125.
- Rothwell PM. Analysis of agreement between measurements of continuous variables: general principles and lessons from studies of imaging of carotid stenosis. *J Neurol* 2000;247:825–834.
- Jobe AH, Bancalari E. Bronchopulmonary dysplasia. *Am J Respir Crit Care Med* 2001;163:1723–1729.
- Beresford MW, Parry H, Shaw NJ. Twelve-month prospective study of oxygen saturation measurements among term and preterm infants. *J Perinatol* 2005;25:30–32.
- Vom Hove M, Prenzel F, Uhlig HH, Robel-Tillig E. Pulmonary outcome in former preterm, very low birth weight children with bronchopulmonary dysplasia: a case-control follow-up at school age. *J Pediatr* 2014;164:40–45.e4.
- Carlo WA. Long-term respiratory morbidities after bronchopulmonary dysplasia despite current therapies. *J Pediatr* 2014;164:12–13.
- Madan A, Brozanski BS, Cole CH, Oden NL, Cohen G, Phelps DL. A pulmonary score for assessing the severity of neonatal chronic lung disease. *Pediatrics* 2005;115:e450–e457.
- Norfolk and Norwich University Hospital. Shunt curve slide rule [accessed 2011 Feb 4]. Available from: <http://www.noranaes.org/shuntcurves/shuntcurves.html>