Inhaled Nitric Oxide Usage in Preterm Infants in the NICHD Neonatal Research Network: Inter-site Variation and Propensity Evaluation

William E. Truog, MD1, Leif D. Nelin, MD2, Abhik Das, PhD3, Douglas E. Kendrick, MS3, Edward F Bell, MD4, Waldemar A. Carlo, MD5, Rosemary D. Higgins, MD6, Abbot R. Laptook, MD7, Pablo J. Sanchez, MD8, Seetha Shankaran, MD9, Barbara J. Stoll, MD10, Krisa P. Van Meurs, MD11, Michele C. Walsh, MD12, and the NICHD Neonatal Research Network

1Children’s Mercy Hospitals and Clinics and the University of Missouri-Kansas City School of Medicine, Kansas City, MO

2Nationwide Children’s Hospital, Columbus, OH

3Statistics & Epidemiology Unit, RTI International, Rockville, MD

4Department of Pediatrics, University of Iowa, Iowa City, IA

5Division of Neonatology, University of Alabama at Birmingham, Birmingham, AL

6Eunice Kennedy Shriver National Institute of Health & Human Development, National Institutes of Health, Bethesda, MD

7Department of Pediatrics, Women & Infants’ Hospital, Brown University School of Medicine, Providence, RI

8Department of Pediatrics, University of Texas Southwestern Medical Center, Dallas, TX

9Department of Pediatrics, Wayne State University School of Medicine, Detroit, MI

10Department of Pediatrics, Emory University School of Medicine & Children’s Healthcare of Atlanta, Atlanta, GA

11Department of Pediatrics, Stanford University School of Medicine, Palo, Alto, CA

12Department of Pediatrics, Rainbow Babies & Children’s Hospital, Case Western Reserve University, Cleveland, OH, for the NICHD Neonatal Research Network

Abstract

Background—The use of inhaled nitric oxide (iNO) in preterm infants remains controversial. In October 2010, an NIH consensus development conference cautioned against use of iNO in preterm infants.
Objective—1) To determine prevalence and variability in use of iNO in the NICHD Neonatal Research Network (NRN) before and after the consensus conference and 2) separately, to examine associations between iNO use and severe BPD or death.

Design/Methods—The NICHD NRN Generic Database collects data including iNO use on very preterm infants. A total of 13 centers contributed data across the time period 2008–2011. Infants exposed or not to iNO were compared using logistic regression, which included factors related to risk as well as their likelihood of being exposed to iNO.

Results—A total of 4,885 infants were assessed between 2008–2011; 128 (2.6%) received iNO before Day 7, 140 (2.9%) between Day 7 and 28 and 47 (1.0%) at >28 days. Center-specific iNO use during 2008–2010 ranged from 21.9% to 0.4%; 12 of 13 sites reduced usage and overall NRN iNO usage decreased from 4.6% to 1.6% (p<0.001) in 2011. Use of iNO started between Day 7 and Day 14 was more prevalent among younger infants with more severe courses in Week 1 and associated with increased risk of severe BPD or death (OR 2.24;95% CI 1.23–4.07).

Conclusions—The variability and total use of iNO decreased in 2011 compared to 2008–2010. iNO administration started at ≥Day 7 was associated with more severe outcomes compared to infants without iNO exposure.

Keywords
Inhaled nitric oxide; bronchopulmonary dysplasia; extremely premature infant

INTRODUCTION

Inhaled nitric oxide (iNO) was approved by the Food and Drug Administration (FDA) in 1999 for use for hypoxic respiratory failure in infants greater than 34 weeks gestational age (GA). The primary outcome of studies evaluating use of iNO in term and near term infants was reduction in the use of extracorporeal membrane oxygenation (ECMO) (1). Based on these encouraging findings and evidence of safety in term infants, several large trials were undertaken to test the benefits, if any, of iNO in very preterm infants. Schreiber (4) and Ballard et al (5) demonstrated improved overall pulmonary outcomes. In contrast, three trials failed to show overall benefit (6–8). However, in post-hoc analyses, two of these trials demonstrated a reduction in BPD in subsets of infants >1000 gm birthweight (6,7). The follow up of the Schreiber et al study (4) also showed improved neurodevelopmental outcomes (9,10). Hibbs et al (11), for the NO-CLD study (5), documented improved pulmonary outcomes based on reduced pulmonary medication use up to 1 year corrected age. In addition, Zupancic et al (12) demonstrated a potential reduction in medical charges in a cost/benefit analysis of iNO for infants who had been enrolled in the second week of life in the Ballard et al (5) study. However, two published meta-analyses (13, 14) concluded that there was no benefit and an individual patient data meta-analysis also concluded that there was no overall benefit to iNO in preterm infants (15). Recently, Clark et al (16) published evidence from the Pediatrix Medical Group administrative database that iNO use had dramatically increased, especially in the most preterm infants, between 2001 and 2008, and was used in 6% of 23–26 week GA infants.
In a recent report, Padula et al (17) demonstrated a 19% rate of iNO usage in infants with severe BPD born at <32 weeks who were treated at 24 large children’s hospitals.

Because of uncertainty regarding the use of iNO in preterm infants, the National Institutes of Health (NIH) convened a consensus development conference in October 2010 (18). The conference’s expert committee concluded that available data did not support routine use of iNO in infants <34 weeks gestation and that additional clinical and translational research should be conducted (18).

The Generic Database (GDB) Registry of the NICHD Neonatal Research Network (NRN) offers an excellent opportunity to achieve two separate goals. Our first goal in this study was to examine the range of iNO use in extremely low GA infants across and within NRN sites, both before and after the NIH consensus development conference statement. Our second goal was to evaluate iNO use and the outcomes of severe BPD or death in very preterm infants who were treated in NRN centers. We used a propensity score modeling approach (19) to first identify the factors associated with iNO use in these infants and to generate probabilities (propensity scores) of iNO use for each infant. We then used these score levels to adjust for any systematic differences in babies treated vs not treated with iNO before comparing the rates of severe BPD or death between these two groups. These analyses incorporated all the infants in both epochs.

**METHODS**

**Neonatal Research Network Generic Database**

The Eunice Kennedy Shriver National Institutes of Child Health and Human Development NRN has maintained a data registry of very low birth weight infants since 1987. This registry has IRB approval at each member site. Since 2008, the registry has collected clinical data on infants who are inborn with a gestational age (GA) >22 weeks and <29 weeks or a birthweight of >400 grams and <1000 grams. Perinatal and routine clinical data and outcomes, including discharge, death, or status at 120 days, are recorded by trained observers and submitted to the Data Coordinating Center (RTI International, Research Triangle Park, North Carolina). GDB infants undergo a BPD classification evaluation at 36 weeks PMA (20, 21). Severe BPD is defined as requiring supplemental O2 for ≥28 days and positive pressure ventilation or FiO2 >0.3 at 36 weeks PMA.

Routine information collected in the GDB includes iNO use and the day of life the iNO administration began. Information regarding initial dosing, duration, escalation, or de-escalation of the dose and indications for starting or stopping is not collected.

The NRN, based on competitive renewal, is funded by the NICHD every 5 years with some previously active sites no longer included and some new sites added with each five-year cycle. To assure a uniform approach across sites, we utilized data for the current analysis derived only from the 13 sites that continuously contributed to the GDB between 2008 and 2011.
To achieve the first goal of this analysis, the overall cohort of infants was divided between two epochs based on birth year: Epoch 1 - calendar years (CY) 2008–2010 and Epoch 2 - CY 2011. These two epochs allowed for comparisons of iNO use before and after recommendations of the NIH consensus conference.

Inclusion criteria for infants for the adjusted analysis examining association between NO use and outcomes using propensity score modeling consisted of all infants who started iNO at ≥7 days or any control infant that never started iNO. Exclusion criteria included infants born with severe congenital anomalies and infants for whom the age of iNO administration could not be identified, and infants treated with iNO at <7 days of age. These time-points for iNO use were based on meta-analyses (13,14) demonstrating that iNO started at ≥7 days would be associated with attempts to reduce BPD.

**Propensity Score Modeling**

Propensity score modeling is a preliminary analysis step that occurs before the actual comparison of the groups of interest, and its purpose is to even out any imbalance between the two groups (to the extent permitted by available data), since iNO treatment was not randomized and there were no controlling schemes that might have checked such imbalances when iNO treatment was initiated (22, 23). A propensity score is created from a logistic regression model which quantifies the probability of receiving iNO treatment, as would be predicted by a set of well-chosen covariates in the propensity model. The propensity score is then used in an analysis model (logistic regression) for comparing the primary outcome (severe BPD or death) between the two treatment groups. Once we obtained the propensity scores, they were ranked and divided into similar groups based on their ranks. Thus, twelve strata were created, each consisting of both treated and control infants that were balanced on key covariates for these two types of infants. These strata were included in the logistic regression analysis model. The modeling approach in this paper can thus be expressed as follows:

**Step 1.** The probability of iNO use was characterized using a set of baseline risk factors considered as covariates or confounders of iNO use in a logistic regression model. The propensity scores (probability of receiving iNO) for each infant were derived from this model. All the data were ranked by propensity score value and divided into twelve groups (propensity score levels).

**Step 2.** The probability of severe BPD or death was characterized using logistic regression incorporating iNO use, propensity score level and other covariates appropriate for that outcome.

**Statistics**

For comparing use of iNO between the two epochs of 2008–2010 and 2011, the χ2 test was used.

The covariates included in the propensity score model (Step 1) for infants starting iNO at ≥7 days included gender, gestational age, small for gestational age (SGA) status at birth, level of FiO2 at Day 7, antenatal steroids, epoch 1 or 2, need for supplemental O2 in the first 7
days, and NRN site. In addition, we investigated adding several other potentially relevant covariates to the propensity score model, including statistical interactions between epoch and other child characteristics (to allow for the possibility that there may be differences in the profile of babies treated with iNO between the two epochs). However, none of these factors approached statistical significance or changed the model in any appreciable manner, and/or were computationally impossible to accommodate given our sample size and the low prevalence of iNO use.

Logistic regression analysis (Step 2) was used for investigating the adjusted association between iNO use and severe BPD or death. Odds ratios and 95% confidence interval were estimated for iNO initiated in the period of Day 7 through Day 14, and also for iNO initiated on or after Day 15. The covariates adjusted for in this model included center, gestational age, birth weight, SGA at birth, male gender, mother’s educational level, antenatal hemorrhage, Apgar score at 5 minutes, need for delivery room resuscitation, prenatal care visit, clinical chorioamnionitis, multiple gestation, major morbidity by Day 7, time group, and propensity score level, indicating likelihood of iNO receipt (obtained from Step 1). A major morbidity by Day 7 included the following: proven sepsis, necrotizing enterocolitis, or Grade ≥3 IVH. The final analysis data set was limited to the 3,582 observations whose propensity scores were within a common range of support. Almost all of the excluded data were from infants who had very low propensity scores, and there were no iNO treated cases with scores that low.

RESULTS

Demographic Data

The population of patients from the NRN GDB utilized for the analysis is shown in Figure 1. Data from a total of 4,979 infants free of identified malformations or syndromes were contributed from the 13 sites. Of those, 83 received iNO but no start date was recorded and 11 infants were missing data on iNO use. A total of 128 infants received iNO in the first 6 days and were not included in the primary analysis (Steps 1 and 2), for a total of 4,757 patients available for analysis (Figure 1).

The 13 centers contributing data represented both larger and smaller centers in the NRN, with a range of 165 to 685 infants submitted to the GDB registry cumulatively over the two epochs. Dividing the data into the two epochs provided the demographic information in Table 1. There were no differences in overall mortality rates between the two epochs (Table 1)

Changes in iNO Use

Initiation of iNO was categorized by day of life; <7 days (128 infants or 2.6%), 7–28 days (140 infants or 2.9%), or >28 days (47 or 0.96%). For iNO use starting on or after Day 7, overall use among all 13 sites decreased from 4.6% in epoch 1 to 1.6% in epoch 2 (p<0.001) (Fig 2). In addition, 12 of the 13 sites demonstrated a significant reduction in epoch 2, often eliminating the use of iNO in this population at these sites. Thus, there was an overall reduction as well as a reduction in inter-site variability. In 2011, only 13 infants in the GDB
were treated with iNO starting between DOL 7–28, and 5 started after 28 days. Infants treated with iNO were smaller and younger than those not treated (693 ± 174 vs. 878 ± 236 gm, and 24.8 ± 1.5 vs. 26.1 ± 1.6 weeks, both p <0.001)

Association with Death or Severe BPD

Propensity score modeling (Step 1) was conducted for the population of infants born between 2008 and 2011 who had initiation of iNO at >7 days or who never started iNO. Use of iNO was associated with lower gestational age and with significant respiratory support by 7 days of age, factors likely associated with worse outcomes (Table 2).

Odds ratios for the adjusted association between first iNO exposure and the primary outcome (severe BPD or death, Step 2) were subsequently determined. Even when the evaluation was limited to initiation of iNO use between 7–14 days, the propensity adjusted analysis demonstrated that use of iNO was associated with increased risk of severe outcome adjusted OR for severe BPD or death =2.24, 95% CI 1.23–4.07.

DISCUSSION

Summary of Results

The current report demonstrates that within the NRN clinical sites, there was substantial variability in the use of iNO for treatment of extremely preterm infants during 2008–2010. Following, and perhaps in response to, the NIH Consensus Development Conference (18), overall use of inhaled nitric oxide decreased dramatically in 2011.

The second and separate result of this study is that the use of iNO in extremely preterm infants in the NRN was not associated with improvement in death or in incidence of severe BPD in infants <29 weeks gestational age. The current finding is consistent with previous meta-analyses (13–15). Based on information available by 2008, infants treated in the first 6 days would likely have been treated with iNO for pulmonary hypertension or severe acute hypoxic respiratory failure to try to improve short-term survival without regard for risk for BPD. However, reasons for initiating iNO in each case were not recorded, nor was duration of treatment, an important factor for determination of justifiable benefit. Use or not of iNO had no overall impact on mortality between the 2 epochs (Table 1), suggesting that its use was limited to already severely ill infants. Indeed, we speculate that many of the infants treated with iNO would not have qualified for the RCTs, based on uncertainty about their long-term survival.

The present report is the first to our knowledge to demonstrate, across a wide spectrum of practices, a reduction in iNO use following the consensus development conference. The wide variability in the use of iNO and conflicting but overall negative findings from meta-analyses and systematic reviews of iNO use in clinical trials in very preterm infants led to the consensus development conference. These meta-analyses (13–15) organized the results either as early use to prevent BPD starting within 2–3 days after birth, selected use of iNO in infants defined as at high-risk for BPD to ameliorate their outcome, or as emergent rescue for critically ill infants.
The present study utilized the propensity score modeling approach to adjust for critical baseline differences between babies who received iNO versus those who did not. The two-step modeling required by this approach not only permits causal inference from observational data, but also provides a natural framework for a systematic examination of the factors associated with receipt of iNO and the adjusted association of iNO treatment with outcomes.

There are several strengths to the current report. The GDB includes infants at sites across a wide geographic spectrum. In addition, sites with diverse populations of extremely preterm infants are included. Based on the comparison of infant characteristics, prenatal referral patterns or prenatal/perinatal practices did not apparently change between the two epochs (Table 1). The database is collected by research coordinators who are skilled in data reliability. Many earlier reports had utilized a dichotomous definition of BPD, which can mask degrees of severity and hence any possible improvement in subsets of infants.

There are several limitations to this study. First, it is difficult to completely account for severity of illness, and it is probable that iNO was utilized in the most severely ill infants, i.e. as emergent rescue. While we used propensity scoring to make the two groups comparable, the approach may still not be able to account for unmeasured or unknown risk factors. Second, for completeness of the dataset, the analysis included infants enrolled through 2011. Because not every infant treated with iNO had a confirmed start date, not every infant treated with iNO was included in these analyses. Third, this report reflects treatments applied to patients born in 13 large perinatal centers. These extremely preterm infants may not be representative of the entire U.S. population of very preterm infants, as 25% of such infants in the United States are born outside large centers (24). Fourth, there was no institutional restriction on the use of iNO at any of the sites, based on protocols in place either before or during 2008–2010. However, external factors, including payment denial by either public or commercial medical insurance providers, may have restricted its use. Such external factors would not be apparent in our dataset. In a report utilizing administrative data from large children’s hospitals, Stenger et al (25) showed that iNO usage rates in a larger group of preterm infants <34 weeks showed a geographic variation. Fifth, we assessed as outcomes only severe BPD or death. It is possible that iNO use may have affected other outcomes, i.e. mild or moderate BPD. Sixth, the data set is limited to infants <29 weeks. It is possible that iNO may benefit somewhat more mature preterm infants who still retain some risk of development of severe BPD. Seventh, these results reflect individual decisions made outside the bounds of rigorous clinical trials. Dose and duration were not recorded and were not available for analysis. Therefore, these data will be excluded from individual patient data (IPD) meta analysis. Some critically ill infants may have been treated with iNO in an effort to stabilize their pulmonary function. If there was no prompt improvement, iNO might have been promptly discontinued.

The results of Ballard et al (5) suggested a longer course of treatment was needed for any effect in BPD. Van Meurs et al (6) found, in a post hoc analysis, improved outcomes in treated infants >1000 gm birthweight; some of those infants were likely born at ≥29 weeks, compared to the treated infants in the current dataset. Also, Schreiber et al (4) showed that, by oxygenation index, less severely ill infants were the responders to iNO. These infants
may have been less likely to be treated in this “field-level” experience. Our findings are consistent with Stenger et al (25), using propensity analysis, that iNO use was associated with increased risk of mortality. Our findings are consistent with the recommendations of the Canadian Pediatric Society Fetus and Newborn Committee (26) and the Committee on Fetus and Newborn of the American Academy of Pediatrics (26) statements on iNO use.

In summary, based upon this comparison of two recent epochs, with apparently similar patients born in 13 large neonatal centers, iNO administration was reduced to a very low frequency in extremely preterm infants following the NIH consensus development conference statement (18). Our study also found no benefit as assessed by death or development of severe BPD from the use of iNO in this population of infants, especially those who were already severely ill at Day 7 of life. The appropriate use of iNO in very preterm infants, if any, remains unclear. However, now there are two studies of preterm infants which used propensity analysis, and neither finds any evidence to support iNO use in infants <29 weeks or <1000 grams. Relative strengths and weaknesses of RCTs and of observational studies utilizing large datasets have been recently examined with advantages of each outlined (28, 29). Completion, analysis, and publication of two additional trials utilizing iNO as part of the protocol design (clintrials.gov NCT 01022580) and (NCT00551642) should help inform future recommendations about iNO use in very preterm infants.

**Abbreviations**

- **iNO**: Inhaled nitric oxide
- **GA**: Gestational age
- **BPD**: Bronchopulmonary dysplasia
- **CY**: Calendar year
- **FDA**: Food and Drug Administration
- **NIH**: National Institutes of Health
- **NICHD**: National Institute of Child Health and Human Development
- **BW**: Birthweight
- **SGA**: Small for gestational age
- **DR**: Delivery Room Resuscitation
- **PPV**: Positive-pressure ventilation
- **OR**: Odds ratio
- **GDB**: Generic Database
- **ECMO**: Extracorporeal membrane oxygenation
References


Figure 1.
Enrollment of GDB patients; of 6,337 patients born between 2008 and 2011 who qualified for the GDB 4,757 were utilized for propensity analysis based on the exclusion criteria shown.
Figure 2. iNO Initiation at ≥7 Days of Age (% of Patients in the Database by Site)
Use of iNO in the 13 NRN sites before (2008–2010) and after (2011) the NIH consensus development conference report. Each individual line represents one NRN site, and iNO use decreased from before to after the NIH consensus development conference report in 12 of the 13 centers. The solid short blue horizontal lines represent mean use of iNO before and after, and use was significantly less after than before the NIH consensus development conference report (p<0.001).
Table 1

Maternal and Infant Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Epoch 1 N=3,635</th>
<th>Epoch 2 N=1,122</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medians</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birthweight (grams)</td>
<td>850 (690, 1,042)</td>
<td>850 (690, 1,030)</td>
<td>0.67</td>
</tr>
<tr>
<td>Gestational age, weeks</td>
<td>26 (25, 27)</td>
<td>26 (25, 27)</td>
<td>0.66</td>
</tr>
<tr>
<td>Apgar score at 5 minutes</td>
<td>7 (5, 8)</td>
<td>7 (5, 8)</td>
<td>0.10</td>
</tr>
<tr>
<td>Age of mother, years</td>
<td>27 (22, 32)</td>
<td>28 (22, 32)</td>
<td>0.05</td>
</tr>
<tr>
<td>Percentages</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small for gestational age</td>
<td>7.5</td>
<td>9.4</td>
<td>0.05</td>
</tr>
<tr>
<td>Male gender</td>
<td>50.8</td>
<td>51.2</td>
<td>0.88</td>
</tr>
<tr>
<td>Resuscitation</td>
<td>85.0</td>
<td>86.7</td>
<td>0.16</td>
</tr>
<tr>
<td>Multiple birth</td>
<td>25.0</td>
<td>24.2</td>
<td>0.58</td>
</tr>
<tr>
<td>Outborn 1</td>
<td>0.2</td>
<td>0.3</td>
<td>0.45</td>
</tr>
<tr>
<td>Race 2</td>
<td></td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>44.3</td>
<td>42.3</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>50.0</td>
<td>52.2</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>5.7</td>
<td>5.5</td>
<td></td>
</tr>
<tr>
<td>C-section</td>
<td>66.3</td>
<td>64.3</td>
<td>0.22</td>
</tr>
<tr>
<td>Overall mortality</td>
<td>27.3%</td>
<td>27.4%</td>
<td>0.90</td>
</tr>
<tr>
<td>Died &lt;12 hours</td>
<td>12.1%</td>
<td>12.0%</td>
<td></td>
</tr>
<tr>
<td>Died &gt;12 hours</td>
<td>15.2%</td>
<td>15.4%</td>
<td></td>
</tr>
</tbody>
</table>

Data are shown as medians (Q1, Q3) or percentages. Medians were tested with a Wilcoxon test. Percentages were tested with a continuity-adjusted chi-square test.

1 Did not meet requirements for chi-square test. Tested with Fisher’s exact test.

2 Relatively large number of missing data. Race was missing for 84 observations; married status was missing for 74 observations.
### Table 2

Propensity score model: Factors associated with iNO at >7 days

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Odds Ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gestational age, weeks</td>
<td>0.68</td>
<td>(0.60, 0.77)</td>
</tr>
<tr>
<td>Small for gestational age</td>
<td>1.59</td>
<td>(0.94, 2.68)</td>
</tr>
<tr>
<td>Male</td>
<td>1.38</td>
<td>(1.00, 1.91)</td>
</tr>
<tr>
<td>Antenatal steroids</td>
<td>1.79</td>
<td>(0.98, 3.28)</td>
</tr>
<tr>
<td>FiO₂ Level(1) at Day 7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal cannula vs room air</td>
<td>2.63</td>
<td>(0.57, 12.13)</td>
</tr>
<tr>
<td>FiO₂ &gt; 0.21 vs room air</td>
<td>3.14</td>
<td>(1.54, 6.39)</td>
</tr>
<tr>
<td>FiO₂ ≥ 0.30 vs room air</td>
<td>7.11</td>
<td>(3.71, 13.62)</td>
</tr>
<tr>
<td>≥1 day of O₂ in first 7 days</td>
<td>0.70</td>
<td>(0.44, 1.10)</td>
</tr>
<tr>
<td>Time group (2011 vs 2008–2010)</td>
<td>0.32</td>
<td>(0.19, 0.53)</td>
</tr>
</tbody>
</table>

Data are shown as odds ratio (OR) and 95% confidence interval (95% CI). Centers were also included in the iNO usage model.

\(1\) Levels were room air without any assistance, nasal cannula but no O₂, FiO₂ > 0.21, and FiO₂ ≥ 0.30.