Commonly Used Neonatal Medications: Practice Patterns Reveal a Need for Comparative Effectiveness Research

Jonathan L. Slaughter, MD, MPH

Nationwide Children’s 2014 Perinatal/Neonatal Conference
May 14, 2014
Practice Variation in Neonatal Medication Prescribing Patterns

• Between Children’s Hospitals
• Between Individual Providers (physicians, nurse practitioners)
• Variation is low when supporting evidence for a treatment is strong (ex. antenatal steroids, surfactant for intubated preterm infants)

Fisher ES, Bynum JP, Skinner JS. NEJM. 2009

• Physician and institutional preference dominate when evidence is lacking
Identifying the Problem

• Lack of evidence upon review of clinical literature

• Widespread variation in practice shows no consensus and the need for further “comparative effectiveness” research
  – Every institution I have been a part of practices neonatology differently (for various treatment decisions: some more intensive/more expensive) and neonatal outcomes “appear” similar
Commonly Used Neonatal Medications with Unproven Effectiveness (and side effects!)

- Chronic Diuretic Use for BPD
- Inhaled steroids
- Bronchodilators (B2-agonists: albuterol, anticholinergic: ipratropium)
- GI meds: proton-pump inhibitors, H2-blockers, motility agents
- Non-steroidal anti-inflammatory drugs
- Antiepileptics
Medication Use in Neonates: extrapolation of adult findings

• Most drugs used in neonates are not FDA approved for neonates, but were instead tested in adults or older children
• Term and especially preterm neonates: have very different and evolving physiologies, changing drug metabolisms
• Immature renal function, chloride-permeable GABA receptor excitatory due to increased intracellular Cl, non-acidic-GERD (spit-up milk) can cause respiratory reactions, etc.
• New drug approvals need to be tested in neonates
• Shouldn’t the testing of existing drugs with unproven effectiveness (and known side effects) be as important a research priority as the development of novel drug therapies?
Long-Term Diuretic Therapy for BPD

- **What’s Known on This Subject:** Diuretics are used in preterm infants to treat the symptoms of bronchopulmonary dysplasia, although there is little evidence of their effectiveness in improving long-term outcomes. Prescribing patterns and frequency of diuretic use in patients with BPD are unknown.

There is little evidence to support any benefit of furosemide administration on need for ventilatory support, length of hospital stay, survival or long-term outcome.

Authors’ conclusions

- In view of the lack of data from randomized trials concerning effects on important clinical outcomes, routine or sustained use of systemic loop diuretics in infants with (or developing) CLD cannot be recommended based on current evidence.
- Randomized trials are needed to assess the effects of furosemide administration on survival, duration of ventilatory support and oxygen administration, length of hospital stay, potential complications and long-term outcome.


Long-Term Diuretic Therapy for BPD

Long-Term Diuretic Therapy for BPD

Percentage of infants by hospital who ever received a >5-consecutive-day course of chlorothiazide, furosemide, hydrochlorothiazide, or spironolactone.

• What This Study Adds: The use of diuretic therapy in infants with BPD, including the specific medications used and length of treatment, varies widely by institution. Long-term diuretic administration to BPD patients is commonly practiced despite minimal evidence regarding effectiveness and safety.

“Despite this lack of evidence to support long-term diuretic use and minimal data on long-term side effects, we found that such use is a routine occurrence in neonatal intensive care units. Surprisingly, a recent survey of 400 U.S. neonatologists by Hagadorn et al to determine factors that influenced a clinician’s decision to use diuretics, found that 66% of respondents expected decreased ventilator days and 59% decreased length of stay.”

Inhaled Nitric Oxide in Preterm Infants

Study data are from 37 children’s hospitals.

Stenger M R et al. Pediatrics 2012;129:e945-e951

©2012 by American Academy of Pediatrics
“Our search identified 557 initial articles and 14 additional studies after reference reviews, with 16 meeting inclusion criteria. Of these, 2 were randomized trials and only 3 additional investigations included comparison groups. We found limited evidence regarding the best pharmacologic treatment for neonatal seizures, but were able to devise a treatment algorithm from available data.”

Main results
In this update we identified four randomised controlled trials investigating the effects of bronchodilators in preterm infants. None of these studies fulfilled our inclusion criterion that clinical outcomes should be reported. One eligible study was previously found dealing with prevention of CLD; this study used salbutamol and enrolled 173 infants. No eligible studies were found dealing with treatment of CLD. Prophylaxis with salbutamol did not show a statistically significant difference in mortality (RR 1.08; 95% CI 0.50 to 2.31; RD 0.01; 95% CI -0.09 to 0.11) or CLD (RR 1.03; 95% CI 0.78 to 1.37; RD 0.02; 95% CI -0.13 to 0.17). No statistically significant differences were seen in other complications associated with CLD or preterm birth. No side effects due to salbutamol were commented on in this study.

Authors’ conclusions
There are insufficient data to reliably assess the use of salbutamol for the prevention of CLD. Further clinical trials are necessary to assess the role of salbutamol or other bronchodilator agents in prophylaxis or treatment of CLD. Researchers studying the effects of bronchodilators in preterm infants should include relevant clinical outcomes in addition to pulmonary mechanical outcomes.

Bronchodilators for BPD

Bronchodilators for BPD

Acid-Suppressive Medications

**Clinical trials have not shown that treating neonatal GERD with H2RAs and PPIs improves GERD symptoms.**


**H2RAs and PPIs are associated with adverse effects that can lead to potentially life-threatening adverse drug events, especially when administered to high-risk neonates.**


Acid-Suppressive Medications

Rates of diagnosis of GERD across participating PHIS hospitals.

Jadcherla S R et al. Hospital Pediatrics 2013;3:335-341

©2013 by American Academy of Pediatrics
## Acid-Suppressive Medications

<table>
<thead>
<tr>
<th>Diagnoses</th>
<th>Ever Received H2RA or PPI</th>
<th>Ever Received H2RA</th>
<th>Ever Received PPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>GERD</td>
<td>7.06* (6.67-7.48)</td>
<td>8.86* (8.29-9.46)</td>
<td>3.85* (3.63-4.08)</td>
</tr>
<tr>
<td>ENT</td>
<td>3.87* (3.54-4.24)</td>
<td>4.59* (4.17-5.06)</td>
<td>2.60* (2.38-2.85)</td>
</tr>
<tr>
<td>GI abnormality</td>
<td>3.34* (3.13-3.55)</td>
<td>2.63* (2.44-2.85)</td>
<td>2.90* (2.72-3.09)</td>
</tr>
<tr>
<td>CDH</td>
<td>3.08* (2.67-3.55)</td>
<td>2.85* (2.44-3.32)</td>
<td>2.36* (2.05-2.71)</td>
</tr>
<tr>
<td>Congenital Heart Disease</td>
<td>2.75* (2.56-2.97)</td>
<td>1.51* (1.37-1.66)</td>
<td>3.00* (2.78-3.24)</td>
</tr>
<tr>
<td>BPD at 36 weeks</td>
<td>1.85* (1.67-2.04)</td>
<td>1.88* (1.65-2.14)</td>
<td>1.66* (1.49-1.85)</td>
</tr>
<tr>
<td>PPHN</td>
<td>1.56* (1.44-1.68)</td>
<td>1.28* (1.15-1.41)</td>
<td>1.59* (1.47-1.73)</td>
</tr>
<tr>
<td>Congenital Lung Disease</td>
<td>1.46* (1.27-1.67)</td>
<td>1.51* (1.28-1.77)</td>
<td>1.44* (1.26-1.66)</td>
</tr>
<tr>
<td>Infection</td>
<td>1.34* (1.27-1.41)</td>
<td>1.29* (1.20-1.38)</td>
<td>1.36* (1.28-1.44)</td>
</tr>
<tr>
<td>Neurologic abnormality</td>
<td>1.19* (1.12-1.27)</td>
<td>1.13* (1.04-1.23)</td>
<td>1.18* (1.11-1.26)</td>
</tr>
</tbody>
</table>

### Birth Gestation

<table>
<thead>
<tr>
<th></th>
<th>≤24 weeks</th>
<th>25-26 weeks</th>
<th>27-28 weeks</th>
<th>29-30 weeks</th>
<th>31-32 weeks</th>
<th>33-34 weeks</th>
<th>35-36 weeks</th>
<th>≥37 weeks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ever Received H2RA</td>
<td>-</td>
<td>0.85* (0.75-0.97)</td>
<td>1.00 (0.85-1.19)</td>
<td>0.85* (0.74-0.97)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ever Received PPI</td>
<td>-</td>
<td>-</td>
<td>0.73* (0.62-0.87)</td>
<td>-</td>
<td>0.62* (0.54-0.71)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

### Gender

<table>
<thead>
<tr>
<th></th>
<th>Ever Received H2RA</th>
<th>Ever Received PPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>0.96 (0.91-1.00)</td>
<td>0.94 (0.88-1.00)</td>
</tr>
</tbody>
</table>

### Intraclass correlation coefficient

<table>
<thead>
<tr>
<th></th>
<th>Ever Received H2RA</th>
<th>Ever Received PPI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.18* (0.12-0.25)</td>
<td>0.29* (0.21-0.39)</td>
</tr>
</tbody>
</table>

Slaughter JL, Reagan PB, Stenger MR, Jadcherla SR. H2RA and PPI Administration to Infants within the NICU. PAS Meeting. 2014
Comparative Effectiveness Research

- **IOM:** "the generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat, and monitor a clinical condition or to improve the delivery of care. The purpose of CER is to assist consumers, clinicians, purchasers, and policy makers to make informed decisions that will improve health care at both the individual and population levels."

- **Federal Coordinating Council (NIH/AHRQ):** *Comparative effectiveness research is the conduct and synthesis of systematic research comparing different interventions and strategies to prevent, diagnose, treat and monitor health conditions. The purpose of this research is to inform patients, providers, and decision-makers, responding to their expressed needs, about which interventions are most effective for which patients under specific circumstances.* To provide this information, comparative effectiveness research must assess a comprehensive array of health-related outcomes for diverse patient populations. Defined interventions compared may include medications, procedures, medical and assistive devices and technologies, behavioral change strategies, and delivery system interventions. This research necessitates the development, expansion, and use of a variety of data sources and methods to assess comparative effectiveness. Systematic research methods can include randomized controlled trials, meta-analyses, observational cohort analyses, and other new and emerging methodologies.

- **Effectiveness in routine-practice is different from efficacy**
Efficacy versus Effectiveness

• **RCTs focused on physiologic outcome:** Do NSAIDs (indomethacin/ibuprofen) close the patent ductus arteriosus?

• **Exclusion of patient from trials:** excluding pregnant women with diabetes from trials of prematurity-related interventions (antenatal steroids)

• “... if recent increases in diabetes incidence continue and diabetes mortality is relatively low, prevalence will increase to 33% by 2050. A middle-ground scenario projects a prevalence of 25% to 28% by 2050.” Boyle *et al.* *Population Health Metrics*. 2010.
“In the patch group, 4 method-failure pregnancies and 1 user-failure pregnancy occurred among 811 women treated for 5240 cycles. In the OC group, 4 method-failure and 3 user-failure pregnancies occurred among 605 women treated for 4167 cycles.”

Patch: 0.06% failure, Pill: 1.2% failure
Effectiveness

- Large prospective cohort study
- 7486 participants
- 334 unintended pregnancies (4.46%)
- IUD/implant: 0.27%
- Pills, patch, ring: 4.55%
- Age <21: doubled risk

Winner et al. Effectiveness of Long-Acting Reversible Contraception. NEJM. 2012

Figure 1. Cumulative Percentage of Participants Who Had a Contraceptive Failure at 1, 2, or 3 Years, According to Contraceptive Method. Bars depict the cumulative percentage of participants who had a contraceptive failure with long-acting reversible contraception (LARC), depot medroxyprogesterone acetate (DMPA), or pill, patch, or ring (PPR) at 1, 2, or 3 years. Participants using PPR had significantly more unintended pregnancies than those using LARC (P<0.001) or DMPA (P<0.001).
Table 3. Hazard Ratio for Unintended Pregnancy, According to Contraceptive Method and Selected Characteristic.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unintended Pregnancy (N = 156)*</th>
<th>Total Participant-Years</th>
<th>Incidence</th>
<th>Hazard Ratio (95% CI)†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>no.</td>
<td>no./100 participant-years</td>
<td></td>
<td>Unadjusted</td>
</tr>
<tr>
<td>Contraceptive method</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LARC</td>
<td>21</td>
<td>7655</td>
<td>0.27</td>
<td>1.00</td>
</tr>
<tr>
<td>DMPA</td>
<td>2</td>
<td>902</td>
<td>0.22</td>
<td>0.72 (0.17–3.09)</td>
</tr>
<tr>
<td>PPR</td>
<td>133</td>
<td>2924</td>
<td>4.55</td>
<td>16.05 (10.19–25.29)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;21 yr</td>
<td></td>
<td></td>
<td></td>
<td>1.83 (1.25–2.69)</td>
</tr>
<tr>
<td>≥21 yr</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Educational level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school or less</td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>Some college</td>
<td></td>
<td></td>
<td></td>
<td>0.73 (0.52–1.03)</td>
</tr>
<tr>
<td>College degree or higher</td>
<td></td>
<td></td>
<td></td>
<td>0.49 (0.28–0.87)</td>
</tr>
<tr>
<td>Previous unintended pregnancies</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>3.22 (1.99–5.19)</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>3.95 (2.29–6.81)</td>
</tr>
<tr>
<td>≥3</td>
<td></td>
<td></td>
<td></td>
<td>5.77 (3.40–9.81)</td>
</tr>
</tbody>
</table>

* The remaining 178 of the 334 unintended pregnancies were attributed to failure of condoms, withdrawal, or any form of contraception that was not included in this analysis. Only periods of index contraceptive-method use (pills, patch, ring, DMPA injection, implant, or intrauterine device) and associated pregnancies were included in this analysis.

† Hazard ratios were adjusted for age, educational level, and number of previous unintended pregnancies.
Patient Data

• Comparative effectiveness experts call for more use of large, observational databases
• Increasingly facilitated by electronic medical records/data collection
• “Real-World” data: better generalizability (external validity)- especially for those who might be excluded from trials
• Better for adverse events (bigger sample size than randomized trials (RCTs))
• Much cheaper and quicker than trials
Figure 1: The Question: What is the Causal Effect of X on Y?

For example:

Membership in the treatment versus control group → Outcome of Interest

\( \beta (?) \)
Figure 2: The Real World

Confounding

• In statistics, a **confounding variable** (also confounding factor, hidden variable, lurking variable, or confounder) is an extraneous variable in a statistical model that correlates (positively or negatively) with both the outcome (dependent) (response) variable and the predictor (independent) (explanatory) variable.

• **Causation**
  – if A causes B:
    • 1) A must precede B
    • 2) change in A must be related/correlated/associated with change in B
    • 3) correlation in A and B is not itself the consequence of correlation between A and B with some prior event C
Propensity scoring: adjusting for observed confounders

• Conditional probability of receiving treatment given the covariates/risk factors; $PS(X) = \Pr(Z=1 \mid X)$

• Obtained by logistic regression. However, instead of the desired outcome, the probability of treatment is the Y (outcome/dependent) variable

• Each patient is assigned a propensity score, a probability (between 0 and 1) of receiving treatment

Glynn, Schneeweiss, and Sturmer. Basic and Clinical Pharmacology and Toxicology. 2006
Propensity Scoring

• Matching- treated patient is matched to an untreated patient with similar propensity score (PS)
• Stratification into Deciles- Treated group and untreated group PS divided into 10 equal groups that are analyzed with each other
• Matching or Stratification:
  • Can be analyzed with simple statistical tests- same as RCT.
  • Separate confounding adjustment from outcome, potentially deters model manipulation until desired outcome (significant result) obtained.
• Inverse Probability of Treatment Weighting of Regression Model
  – Can give Average Treatment Effect (ATE) as well at ATT (Average Treatment Effect on the Treated)
Propensity Scoring: when is it better than multivariable regression?

• Small numbers of events relative to the number of potential confounders
  – Particularly relevant in pharmacoepidemiology: rare outcomes in patients with multiple risk factors
  – When fewer than 8 events per confounder PS estimates less biased, more robust, more precise than logistic regression approach

• When treatment effect may vary by strength of indication for its use (can detect with stratification)

• When a smaller study/database has detailed information/covariates that a larger study/database lack - can use “propensity score calibration”

• When there is non-overlap of the exposure among treated/untreated

Glynn, Schneeweiss, and Sturmer. Basic and Clinical Pharmacology and Toxicology. 2006
Indications for Propensity Scores and Review of their Use in Pharmacoepidemiology
The horizontal axis represents deciles of the sample, grouped according to increasing probability that the patients would receive iNO (ie, the estimated propensity score).

Stenger M R et al. Pediatrics 2012;129:e945-e951
Econometric Techniques

• How might we control for unobserved confounders?

• Instrument: a variable that only effects the outcome through its association with the treatment

• Assumptions necessary (as with all statistics), assumption that a variable is an instrument is hard to test/prove
Instrumental Variables

Z (assignment to vitamin A) → X (receipt of vitamin A) → Y (death during follow-up)

U (unmeasured confounders) → Y (death during follow-up)

Natural Experiment

• A **natural experiment** is an empirical study in which the experimental conditions (i.e., which units receive which treatment) are determined by nature or by other factors out of the control of the experimenters and yet the treatment assignment process is arguably exogenous (outside of the model-instrument).

• Vietnam draft
• Distance to hospital with cath lab
• Drug Shortages
• John Snow (1854)- distance to pump

• Plan to use recent indomethacin/ibuprofen shortages and recalls as a natural experiment in an instrumental variable analysis
Grouped Treatment Variable Analysis

- Patient Level (individual-level) patient covariates/risk factors (some unmeasured) influence clinician decision to treat that patient
- Group level (ecologic) treatment frequency (at the hospital-level) is less sensitive to unmeasured confounders at the patient level
- Exploits Practice Variation
  - Assumptions similar to Instrumental Variable: center/hospital level frequency of treatment independent of outcome except through the treatment
- From our studies of multiple medications, it appears that physician preference/hospital precedent play the largest role in whether an infant is treated
- Grouped treatment variable (hospital-level frequency of tx) can replace individual-level treatment in a regression model, or be used in a two-stage instrumental variable design
- Outcome measurement is also at the group-level: do hospitals that treat more PDAs with NSAIDS, on average, have better respiratory outcomes?

Non-steroidal anti-inflammatory drugs (NSAIDs) have been widely used for over three decades to close the patent ductus arteriosus (PDA) in preterm infants. However, it is unclear whether treatment closure of PDA improves patient outcomes. The ductus, an essential component of the fetal circulation, normally closes shortly after birth in term infants. However, in 65% of preterm infants born prior to 28 weeks ductal patency persists. The presence of PDA in these infants is associated with increased mortality and higher rates of bronchopulmonary dysplasia (BPD). Ductal closure aims to reduce these complications, but results of recent observational studies have failed to show improved outcomes despite successful ductal closure and even raised the possibility of harm following treatment. NSAIDs decrease cerebral, mesenteric, and renal blood flow in infants and acute renal insufficiency is common. Indomethacin has been associated with intestinal perforation and maternal administration was recently associated with neonatal necrotizing enterocolitis. Furthermore, NSAID therapy in neonates is expensive and prone to frequent shortages. When left untreated, the majority of untreated PDAs in preterm infants close spontaneously by 44 weeks postmenstrual age. Randomized controlled trials (RCTs) have determined the efficacy of NSAID treatment to close PDAs, but to date no large treatment trials have focused primarily on associated morbidities. Therefore, there is a critical need to identify whether attempted PDA closure with NSAIDS is warranted in preterm infants.
Figure 3. Variation in the % of infants at each hospital receiving NSAIDs during the first 14 days of age.
A. Indomethacin Prophylaxis  
B. Indomethacin Treatment  
C. Ibuprofen Treatment

Figure 3 includes the 41 hospitals with at least 20 patients meeting inclusion criteria. A. Percentage of infants receiving prophylactic indomethacin in the first 24 hours postnatally (range: 0-35%); B. Percentage receiving treatment with indomethacin beginning after 24 hours but prior to 14 postnatal days (0-34%); C. Percentage receiving treatment with ibuprofen in first 14 postnatal days (0-21%). We also repeated the analyses by including infants treated with NSAIDs up to postnatal day 28. These results were similar with nearly identical graph results showing that most NSAID administration occurs prior to day 14. This was expected since NSAIDs are most effective for PDA closure in the first two postnatal weeks.
Aim 2: To determine the comparative effectiveness of NSAID treatment versus no treatment for PDA on mortality and respiratory outcomes. We hypothesize that NSAID treatment provides no additional improvement in these outcomes, as compared to conservative management. To accomplish this aim, we will employ novel approaches to reduce confounding by indication. Propensity scoring will be used to reduce bias due to measured confounders. Recent indomethacin and ibuprofen supply shortages provide a natural experiment, and the timing of these shortages will serve as instrumental variables to reduce the effect of unmeasured confounders. We will also utilize variation in the hospital-specific frequency of NSAID treatment for PDA (grouped-treatment variable analysis), as a second means to reduce unmeasured confounding. We will determine the effect of NSAID treatment on mortality, duration of mechanical ventilation, duration of oxygen administration, and BPD at 36 weeks postmenstrual age.