Choosing Wisely in Newborn Medicine: Five Opportunities to Increase Value

Timmy Ho, MD,a,b,c,d Dmitry Dukhovny, MD, MPH,a,e John A.F. Zupancic, MD, ScD,a,b,d Don A. Goldmann, MD,a,b,c,d Jeffrey D. Horbar, MD,f,g DeWayne M. Pursley, MD, MPH,a,b,d

abstract

BACKGROUND: The use of unnecessary tests and treatments contributes to health care waste. The “Choosing Wisely” campaign charges medical societies with identifying such items. This report describes the identification of 5 tests and treatments in newborn medicine.

METHODS: A national survey identified candidate tests and treatments. An expert panel of 51 individuals representing 28 perinatal care organizations narrowed the list over 3 rounds of a modified Delphi process. In the final round, the panel was provided with Grading of Recommendation, Assessment, Development and Evaluation (GRADE) literature summaries of the top 12 tests and treatments.

RESULTS: A total of 1648 candidate tests and 1222 treatments were suggested by 1047 survey respondents. After 3 Delphi rounds, the expert panel achieved consensus on the following top 5 items: (1) avoid routine use of antireflux medications for treatment of symptomatic gastroesophageal reflux disease or for treatment of apnea and desaturation in preterm infants, (2) avoid routine continuation of antibiotic therapy beyond 48 hours for initially asymptomatic infants without evidence of bacterial infection, (3) avoid routine use of pneumograms for predischarge assessment of ongoing and/or prolonged apnea of prematurity, (4) avoid routine daily chest radiographs without an indication for intubated infants, and (5) avoid routine screening term-equivalent or discharge brain MRIs in preterm infants.

CONCLUSIONS: The Choosing Wisely Top Five for newborn medicine highlights tests and treatments that cannot be adequately justified on the basis of efficacy, safety, or cost. This list serves as a starting point for quality improvement efforts to optimize both clinical outcomes and resource utilization in newborn care.

Overuse and waste remain significant problems in the US health care system, by one estimate accounting for ~34% of all health care spending in 2011, then assessed at ~$2.7 trillion. In this study, 5 categories of waste were described: failures of care delivery, failures of care coordination, overtreatment, administrative complexity, and pricing failures. Overtreatment ("waste that comes from subjecting patients to care that, according to sound science and the patient’s own preferences, cannot help them") alone was estimated to contribute between $158 billion and $226 billion in wasteful spending in 2011. The reduction or elimination of low-value tests and treatments would reduce costs and could potentially improve quality.1

In 2011, the American Board of Internal Medicine Foundation launched the “Choosing Wisely” campaign, which encourages "physicians, patients, and other healthcare stakeholders to think
and talk about medical tests and procedures that may be unnecessary.2 The foundation charges medical societies with the generation of Top Five lists of tests and treatments to promote “care that is supported by evidence, not duplicative of other tests or procedures already received, free from harm, and truly necessary.”3 The American Academy of Pediatrics (AAP) joined the Choosing Wisely campaign in 2013, but there has been minimal participation by groups representing subspecialty pediatric care.4 To help ensure that care of inpatient newborns is efficient, as well as evidence-based, the AAP Section on Perinatal Pediatrics undertook an extensive process, including surveys, consensus-building methods, and literature reviews, to develop a Top Five list in newborn medicine.

**METHODS**

Figure 1 depicts the flow diagram of the methodology for the development of the AAP newborn medicine Top Five list. The Beth Israel Deaconess Medical Center Committee on Clinical Investigation approved the study as exempt research.

**National Survey of Stakeholders**

A survey was developed to which respondents were asked to consider the full range of tests and treatments conducted on both low- and high-risk newborns. Participants provided at least 1 example and as many as 10 examples of tests and treatments that, in their opinion, best met any or all of the following criteria: (1) evidence of lack of efficacy, (2) insufficient evidence of efficacy, or (3) unnecessary utilization of staffing or material resources. We pretested the instrument for face validity in a group of 10 neonatologists and revised it iteratively. We then administered the survey to the 2872 neonatologists in the AAP Section on Perinatal Pediatrics and to 1053 physicians, nurses, NICU staff, and family representatives attending the Vermont Oxford Network Annual Congress.

**Expert Panel and the Modified Delphi Process**

We assembled an expert panel of 51 individuals representing the leadership of 28 national and regional stakeholder perinatal care groups, including the Executive Committee of the AAP Section on Perinatal Pediatrics, the AAP Committee on the Fetus and Newborn, quality improvement organizations such as the Vermont Oxford Network and state neonatology quality collaboratives, biomedical journals,
and other professional organizations. The expert panel participated in 3 rounds of a modified Delphi process (also known as the Rand/UCLA appropriateness method). This method gathers a group of experts and attempts to achieve agreement through repeated questioning, alternating with controlled opinion feedback. In this study, our modifications included the use of electronic surveys, allowing feedback of only numerical results without the opportunity for face-to-face discussion. In addition, to quantify the strength of opinion, each panel member was given 100 points to distribute among candidate tests and treatments. We asked members to allocate points to items according to how strongly they felt that routine use of that test or treatment met the evidence and utilization criteria described above. Four authors (T.H., J.A.F.Z., D.D., and D.M.P.) acted as moderators, collating deidentified results and presenting the results of each round back to the expert panel. After each round, tests and treatments proceeded to the next round if they achieved both priority (in the top 75% of the item list when ranked by points) and consensus (defined a priori for this purpose as at least 50% of the responding expert panel members assigned points to the item). The use of the consensus criterion ensured that individuals could not distort the results by applying a very high value to a single item.

**Rounds 1 and 2 of the Modified Delphi Process**

We first organized the items suggested by survey participants into general test and treatment categories. We identified and presented to the panel items that had received at least 30 mentions in the survey. This threshold was selected on the basis of the consensus of the investigators that it represented a sufficient number to reflect opinion among respondents while ensuring a feasible number of items for subsequent literature review and expert panel consideration. The expert panel then allocated points on these first-round items. For this round only, tests and treatments that did not achieve both consensus and priority in the weighting process were returned to the expert panel in a subround and assessed with a third criterion: items selected by >10% of the responding panel members were retained for further consideration. We then reintroduced specific clinical indications drawn from the initial survey responses to the resulting tests and treatments before the expert panel again reduced the list during the second round.

**GRADE Literature Review and the Final Round**

For the items remaining after the second Delphi round, 4 authors (T.H., D.D., J.A.F.Z., and D.M.P.) conducted a review of the literature in PubMed and the Cochrane Library to generate Grading of Recommendation, Assessment, Development, and Evaluation (GRADE) Summary of Findings tables. GRADE Summary of Findings tables present results of a literature review in an easy-to-understand table that includes identification of the patient or population, setting, intervention, comparison, outcomes of interest, compiled results from different studies, and a grading of the quality of the evidence. Reviewers first searched for published systematic reviews, including those from the Cochrane Neonatal Review Group. If existing reviews were available, we used the same search parameters to obtain updated citations; otherwise, we conducted de novo independent literature searches. The search included all literature indexed on Medline with publication dates through 2013. We selected studies for reviews on the basis of their adherence to the population, setting, intervention, and comparison of interest for each item. After the GRADE Summary of Findings tables, the results were further distilled into 1-paragraph summaries for review by the expert panel.

In a final round, we presented 1-paragraph summaries (see Supplemental Information) and the GRADE Summary of Findings tables to the expert panel for review. The panel again distributed 100 points among the items according to their estimates of the degree to which items met the Choosing Wisely criteria. The 5 items with the highest total points were included in the newborn medicine Top Five list. The final list was reviewed and approved by the AAP Board of Directors and Executive Committee.

**RESULTS**

We received responses from 1047 individuals, who submitted a total of 2870 suggestions for candidate tests and treatments to be further considered. Demographic characteristics of the respondents are provided in Table 1.

Of the 2870 suggestions, 1648 (57%) were classified as tests, and the remainder were treatments. Among tests, 71% were screening tests, 23%...
were diagnostic tests, and 6% were monitoring studies. Among the screening tests, most were laboratory-based, followed by imaging, bedside, car seat, and congenital heart disease screening oximetry. In the monitoring category, pneumograms were most commonly cited, followed by NICU apnea and oximetry monitoring, and then home monitoring. The great majority of diagnostic testing was imaging, which was followed by laboratory testing. Medications constituted 56% of suggested overused treatments, followed by respiratory procedures (11%), surgical procedures (10%), and nutritional interventions (9%). In the medication category, the top 4 groups were medications for gastroesophageal reflux, followed by antimicrobial agents, diuretics, and patent ductus arteriosus prophylaxis and treatment.

The initial list of 22 tests and treatments was reduced to 13 during the first round. The addition of specific clinical indications increased this number to 24. These were then reduced to 14 tests and treatments with specific clinical indications during the second round. Two similar items were combined (antireflux medications for the treatment of apnea and desaturation and antireflux medications for the treatment of reflux), whereas an item outside of the influence of providers (newborn screening after multiple normal screens is generally mandated by regulatory authorities) was eliminated.

The initial list of 22 tests and treatments, the second list of 24 tests and treatments with specific clinical context, and the penultimate list of 12 tests and treatments that underwent literature review can be found as electronic supplements to this publication (see Supplemental Tables 3, 4, and 5).

The newborn medicine Top Five list generated in the final Delphi round is shown in Table 2. The Top Five list is further described below with a 1-paragraph summary of the rationale targeted to the general audience. Note that the wording of the following items was both drawn directly from survey suggestions and presented to the expert panel without edits or elaboration from the research group.

### Avoid Routine Use of Antireflux Medications for Treatment of Symptomatic Gastroesophageal Reflux Disease or for Treatment of Apnea and Desaturation in Preterm Infants

Gastroesophageal reflux is normal in infants. There is minimal evidence that reflux causes apnea and desaturation. Similarly, there is little scientific support for the use of H2 antagonists, proton-pump inhibitors, and motility agents for the treatment of symptomatic reflux.6,7 Importantly, several studies show that their use may have adverse physiologic effects as well as an association with necrotizing enterocolitis,8 infection,9 and possibly intraventricular hemorrhage and mortality.10

### Avoid Routine Continuation of Antibiotic Therapy Beyond 48 Hours for Initially Asymptomatic Infants Without Evidence of Bacterial Infection

There is insufficient evidence to support antibiotic treatment of >48 hours to rule out bacterial infection in asymptomatic preterm infants. Current blood-culturing systems identify the great majority of pathologic organisms before 48 hours.11 Prolonged antibiotic use may be associated with necrotizing enterocolitis and death in extremely low birth weight infants.12

### Avoid Routine Use of Pneumograms for Predischarge Assessment of Ongoing and/or Prolonged Apnea of Prematurity

Cardiorespiratory events are common in both term and preterm infants.13 Although there may be a role for pneumograms in selected cases in which the etiology of the events is in doubt, they have not been shown to reduce acute life-threatening events or mortality from their routine use.14

### Avoid Routine Daily Chest Radiographs Without an Indication for Intubated Infants

Although intermittent chest radiographs may identify unexpected findings, there is no evidence documenting the effectiveness of daily chest radiographs to reduce adverse outcomes. Furthermore, this practice is associated with increased radiation exposure.15,16

### Avoid Routine Screening Term-Equivalent or Discharge Brain MRIs in Preterm Infants

Findings on term-equivalent MRI correlate with neurodevelopmental outcomes at discharge and at 2 and 5 years of age.17 There is, however, insufficient evidence that the routine use of term-equivalent or discharge screening brain MRIs in preterm infants improves long-term outcome.18–24

### DISCUSSION

We present the results of a 2-year process, involving a broad coalition of...
stakeholders and supported by systematic literature review, to promote evidence-based, value-conscious care in newborn medicine. Together, we identified 5 areas of overuse, including routine uses of antireflux medications, antibiotic therapy beyond 48 hours, pneumograms, daily chest radiographs, and term-equivalent brain MRIs. The literature shows evidence of harm for 2 of these 5 items (antireflux medications and antibiotic therapy beyond 48 hours), whereas the remaining 3 items (pneumograms, daily chest radiographs, and term-equivalent brain MRIs) lack sufficient evidence of efficacy. This latter group will benefit from further study.

To date, 63 medical societies in the United States and several organizations in at least 12 countries have created Top Five lists in the Choosing Wisely campaign.25 Although the majority of lists address issues in adult medicine, the AAP, the Society of Hospital Medicine (represented by its Pediatric Choosing Wisely Committee), and the American College of Rheumatology (Pediatric Rheumatology) have also participated.26,27 Newborn medicine offers special challenges to the Choosing Wisely approach. Neonatal care is provided by an interprofessional team through which a family's influence has been somewhat indirect. Parents of neonates, with help from medical providers, determine long-term care goals, but providers could offer more opportunities for active participation in daily clinical decisions including routine screening and diagnostic tests or noninvasive treatments. This issue represents an area of opportunity for improvement in newborn medicine; the Top Five list can serve as a catalyst for more shared decision-making. Indeed, even without a Choosing Wisely Top Five list, NICUs in other countries report greater parental involvement in the care of preterm infants.28,29 To ensure the relevance of Choosing Wisely to this greater involvement, we included parents, as well as other representatives of the neonatal care team, in both the survey component and on the expert panel of this process. Another distinctive characteristic of this Top Five list is that neonatal care is largely inpatient-focused. Hospitals will be increasingly challenged to provide efficient care as payment evolves from fee-for-service to case-based to population-based reimbursement models.30 At the same time, however, hospitals currently engaged in robust quality improvement efforts may be better equipped to leverage culture, knowledge, and tools to meet the challenges of the Choosing Wisely campaign.

Unlike purely consensus-driven guidelines, our study benefits from multiple methods that ensure both internal validity and generalizability. The initial survey was administered to a multidisciplinary group of >1000 frontline practitioners and parents. As with the Top Five list from the Pediatric Choosing Wisely Committee of the Society of Hospital Medicine, the newborn medicine list also uses the Rand/UCLA appropriateness method, or modified Delphi method, to build consensus.31 This work builds on that strength by involving an expert panel with diverse representation.32 Last, a thorough review of the literature and distillation of that knowledge using the GRADE-validated approach to evidence synthesis helped to ensure that all current evidence was considered.

Interestingly, with the exception of term-equivalent brain MRIs, most of the items on the list are not associated with high prices. However, even low-priced tests and treatments can have significant cost implications due to volume, use of staff time, or triggering of subsequent testing and therapies and might therefore represent appropriate targets for a value-focused approach.33,34 Consider that the elimination of even 1 radiograph, assuming even a conservative price of $100 per film, for each very low birth weight infant in an NICU with 200 such admissions per year would represent savings of >$20 000. Moreover, the elimination of multiple low-price ineffective therapies might cumulatively impact efficiency at a national level.

Variations in practice in NICUs have been well established.35–37 Similarly, the incidence of items on the Top Five list will vary across centers. In applying the results, we would encourage centers to first measure incidence of use. Certain items may be prioritized for reduction or elimination. Once high reliability has been achieved in the top 5 items, such self-examination will reveal other areas of overuse and waste on which value improvement efforts ought to focus. For example, a National Institutes of Health consensus panel has recommended against the administration of nitric oxide to most preterm infants,38 but this practice is not included in the Top Five list. Nitric oxide use may not be an issue in the large majority of centers, but reduction in its overuse will have a substantial cost impact in certain outliers.

Readers will note the repeated reference to “routine” in the Top Five list. It is important to acknowledge that there may be specific circumstances in which these tests and treatments may be appropriate or even indicated. Standardization of “best practice” never should override clinician judgment based on the newborn's clinical presentation. However, as part of a learning health care system, if tests and treatments on the Top Five list are ordered, the clinician should specify why the order was justified. Variance from standard practice should be systematically examined to ascertain under which circumstances deviation may be justifiable or preferred. Certainly the
lack of evidence for some of these practices should not inhibit their inclusion in well-designed research protocols. For example, given the reported predictive validity of term-equivalent MRI, it will also be critical to establish that it results in improved outcomes and does so more effectively than less costly modalities. Accordingly, policy makers and payors should not interpret this Top Five list and other Choosing Wisely Top Five lists as “Do Not Do” lists used to justify penalties and censure, rather than improvement and learning.

This Top Five list serves as a starting point for the continuous process of increasing value in care delivery. Additional research and inquiry may reveal an evolving set of tests or treatments that are not efficacious and incur unwarranted costs. In fact, this research group hopes that as overuse of certain low-value tests and treatments decreases and new knowledge about efficacy comes to light, other tests and treatments will rise to the forefront of a new value-focused approach to newborn care. Local priorities and circumstances will influence which aspects of overuse are tackled first, and respecting staff perspectives and recommendations is important in catalyzing improvement. Ideally, clinical units will also go beyond a strict focus on overuse and would include efficacious practices that are underused or inefficiently delivered. Local institutions may address these deficiencies by using established quality improvement methods, such as the Model for Improvement, or waste elimination techniques, such as Lean. These approaches may require innovative tools, as well as multidisciplinary education focused on value in neonatal care.

**CONCLUSIONS**

There are reasons to anticipate that waste reduction will be particularly successful in newborn medicine. A team-based care structure ensures multidisciplinary representation in both identifying opportunities for waste reduction and in implementing solutions. Although our results suggest that non-evidence-based practices persist, Neonatology was one of the earliest proponents of evidence-based medicine; there is therefore a substantial literature on which to make recommendations. Highly functioning quality collaboratives at both the state and national levels provide a foundation for measurement and large-scale change. Clinical care guidance and policy are centrally coordinated through the AAP committee structure (http://pediatrics.aappublications.org/site/aappolicy/index.xhtml). The evolving health care delivery system (more transparent quality reporting, the electronic health record, and changing health care payment methods) serves as catalyst for the provision of more efficient care. These 5 assets for success, and value initiatives such as that described in this article, should help ensure that we choose wisely for the sickest, smallest, and most vulnerable patients in our society.

**ACKNOWLEDGMENTS**

We acknowledge the contributions of the survey respondents and of the members of the expert panel including the following: Dr Susan Alcott, Dr Wanda Barfield, Dr William Benitz, Dr Carl Bose, Dr David Burchfield, Ms Madge Buss-Frank, Dr Cheryl Ann Carlson, Ms Joanne Celenza, Dr Morris Cohen, Dr James Cummings, Dr Kara Driver, Dr Eric Eichenwald, Dr William Engle, Dr Ivan Frantz, Dr Thomas George, Dr Mitchell Goldstein, Dr Sergio Golombek, Dr Jeffrey Gould, Dr Peter Grubb, Dr Munish Gupta, Dr Douglas Hardy, Dr Andrew Hopper, Dr Mark Hudak, Dr Michael Kuzniewicz, Dr Edward Lawson, Dr Krithika Lingappan, Dr Lily Lou, Dr Martin McCaffrey, Dr Mary Nock, Dr Akihiko Noguchi, Dr Alfonso Pantoja, Dr Alan Picarillo, Dr Brenda Poin Dexter, Dr Richard Polin, Dr Renate Savich, Dr Roger Soll, Dr Alan Spitzer, Dr Suzanne Staebler, Dr Dan Stewart, Dr Jonathan Swanson, Dr Rosemarie Tan, Ms Sarah Walder, Dr Michele Walsh, Dr Kristi Watterberg, and Dr Cherrie Welch.

We also appreciate the ongoing feedback of the study provided by the fellows and faculty of the Harvard Pediatric Health Services Research Fellowship and the fellows of the Institute for Healthcare Improvement. Thoughtful review of the manuscript was provided by Dr Jonathan Finkelstein, MD, MPH.

**ABBREVIATIONS**

AAP: American Academy of Pediatrics
GRADE: Grading of Recommendation, Assessment, Development and Evaluation

---

**FINANCIAL DISCLOSURE:** Dr Horbar is the Chief Executive Officer of the Vermont Oxford Network (VON); and Drs Ho, Dukhovny, Zupancic, Goldmann, and Pursley have received honoraria for faculty participation from VON.

**FUNDING:** This project was completed without specific support. Dr Ho’s work was supported by an Agency for Healthcare Research and Quality National Research Service Award institutional training grant (5T32HS000683-21).

**POTENTIAL CONFLICT OF INTEREST:** The authors have indicated they have no potential conflicts of interest to disclose.
REFERENCES


Choosing Wisely in Newborn Medicine: Five Opportunities to Increase Value
Timmy Ho, Dmitry Dukhovny, John A.F. Zupancic, Don A. Goldmann, Jeffrey D. Horbar and DeWayne M. Pursley

Pediatrics 2015;136:e482; originally published online July 20, 2015;
DOI: 10.1542/peds.2015-0737

The online version of this article, along with updated information and services, is located on the World Wide Web at:
/content/136/2/e482.full.html