Title: Practice variation may be evidence-based: A multi-criteria decision analysis of treatments of patent ductus arteriosus

Running title: Patent ductus arteriosus decision analysis

Timothy Disher, PhD(c), BScN, RN, Dalhousie University School of Nursing, Halifax, NS, Canada

Souvik Mitra, MD, RCPC Affialite, IWK Health Centre and Dalhousie University, Halifax, NS, Canada

Louis Beaubien, PhD, Dalhousie University, Halifax, NS, Canada

Chris Cameron, MSc, PhD, Cornerstone Research Group, Burlington, ON, Canada

Marsha Campbell-Yeo, PhD, BC-NNP, Dalhousie University School of Nursing and IWK Health Centre, Halifax, NS, Canada

**Conflicts of Interest:** Dr. Cameron is an employee and holds shares of Cornerstone Research Group who provide consultant services to various pharmaceutical and device companies. Mr. Disher is an employee of Cornerstone Research Group. No other authors have conflicts of interest to disclose.

**Funding:** No study-specific funding to report. Disher is supported by a Canadian Graduate Scholarships Vanier Scholarship, a Killam pre-doctoral scholarship, an NSHRF Scotia Scholar award, a Nova Scotia Graduate Scholarship, and the Electa MacLennan memorial scholarship. Campbell-Yeo is supported by a Canadian Institutes of Health Research (CIHR) New Investigator Award.

**Word Count**

**Abstract**

Aim

Methods

Results

Conclusion

**Key notes**

Please sum up your article in three bulleted short sentences. Maximum of 70 words total with the aim of creating an easily digestible take-home message. No introduction of facts not contained in the abstract.

**Keywords**

**Background**

Patent Ductus Arteriosus is a common cardiovascular complication in preterm infants. The resulting hemodynamic consequences of PDA can lead to increased risk for major neonatal morbidities including bronchopulmonary dysplasia, necrotizing enterocolitis, intraventricular hemorrghage, and death. Clinicians weigh these potential outcomes against the possibility of harm associated with pharmacological or surgical closure of the ductus. Current neonatal practice has moved away form early trends of proactive surgical closure to favour conservative and pharmacological approaches. Despite this general consensus move away from surgical management there is significant centre to centre and clinician to clinician variability in choice of treatment. A systematic review and network meta-analysis of hemodynamically significant PDA categorized pharmacological treatments into nine categories evaluated across eight outcomes (PDA closure, need for repeat pharmacotherapy, need for surgical closure, mortality, necrotizing enterocolitis, bronchopulmonary dysplasia, intraventrichular hemorrhage, and oliguria). Interestingly, the relative ranking of treatments varied across these outcomes and thus it would be reasonable to expect continued practice variation arising from differing absolute event rates across centres and relative importance of outcomes across clinicians. Network meta-analyses represent the state of the art in meta-synthesis, their utility for guiding clinical decision making is limited by a lack of formal comparison of the absolute differences in outcomes considered in the context of their relative importance for decision making. Multi-criteria decision analysis is an approach to decision making when evaluations are required across multiple outcomes that differ in perceived importance. Stochastic Multi-Criteria Acceptability Analysis (SMAA) is an approach to MCDA that allows for a formal assessment of treatment alternatives that simultaneously incorporates uncertainty in absolute outcomes as well as uncertainty in the relative importance of outcomes. The purpose of this decision analysis is to assess how optimal treatment decisions vary across baseline event rates and outcome preferences.

**Method**

**Study Design:** We conducted a decision analysis of treatment for hemodynamically significant PDA using stochastic multicriteria decision analysis driven by results from a Bayesian NMA.

**Data sources:** A detailed summary of the main review and NMA methods have been published in detail previously. In summary, eligible trials included randomized controlled trials of neonates born less than 37 weeks gestational age or 2500g who were provided with pharmacological treatment for a hemodynamically significant PDA. PDA could be diagnosed clinically or via echo-cardiograph. The full set of outcomes can be found in table one. A search of MEDLINE, Embase, and Cochrane CENTRAL included articles published up to December 31, 2017 and was supplemented with hand search, trial registry searches, and personal communication. In total, sixty seven articles were included in the network meta-analysis, although total numbers varied by outcome (e.g. 60 trials for PDA closure, 33 trials for repeat pharmacotherapy, eFigure 1). In the original article a sub-group analysis was conducted based of trials considered at low risk of overall bias but results were not meaningfully different and we use the full primary analysis here.

**Statistical Analysis**

**Network meta-analysis.** Bayesian network meta-analysis is ideally suited to decision analysis since it allows for integration of the decision across the entire posterior, without the need reliance on point estimates or assumptions that uncertainty is normally distributed. The primary author of the original publication is a co-author on this project and provided the full set of original data for re-analysis. Consistent with the original analysis, we fit a random effects model using previously published code that places vague priors on treatment effects. We used informative priors on the between trial heterogeneity published by Turner to allow reasonable estimates in the presence of many single study connections. Our analysis implementation differs from the original publication in one significant way, in that we fit a single unified model across outcomes. The reason for this change is that is that some treatments had no information on some outcomes (e.g. high dose oral ibuprofen did not have an estimate for BPD), which makes MCDA impossible. To address this challenge, we adapted approaches originally created for analysis of longitudinal data in which we assume that treatment efficacy is exchangeable across outcomes by adding an additional level to the hierarchical model (e.g. if a treatment reduces NEC it will have similar relative effect on BPD). Within the Bayesian framework, this allows for estimation of treatment effects otherwise not reported by imputing them from the posterior predictive distribution through a concept known as borrowing strength. The use of a suitably vague prior on the standard deviation of this overarching distribution allows for the amount of borrowing of strength to be driven by the data. If the observed treatment effects for an intervention are not consistent across outcomes this standard deviation will increase, the model will borrow less strength across outcomes, and the uncertainty in imputed values will increase. This process introduces a unit of analysis issue in that a single trial can be double counted across outcomes. To account for this, we follow methods published by Jansen in which the variance is inflated by dividing it by 1 – p^2, where p is an assumed correlation across outcomes that we set to 0.5. All outcomes were scaled so that larger odds ratios were associated a positive treatment effect. We ran the NMA model using Just Another Gibbs Sampler (JAGS) on three chains for 100,000 iterations each including a 30,000 interation burn in period. We checked for convergence using the Gelman-Rubin r hat statistic where values less than 1.1 were considered evidence of convergence if traceplots also appeared reasonable.

**SMAA.** Stochastic multicriteria acceptability assessment was conducted using methods outlined by van valkenhoef. First, odds ratios from the NMA were converted to absolute probabilities by assuming that the placebo arm risk was an acceptable estimate of baseline risk. Next, we elicited ordinal outcome rankings from two clinicians (MCY/ SM, Table 2). We used the hitandrun package to simulate from the feasible weight space given these ordinal constraints using a markov chain monte carlo approach. This allows for consideration of the uncertainty of the exact vector of outcome weights given a clinician’s ability to specify that e.g. mortality is more important than NEC without the need to specify an exact relative weight. We supplement this set of weights with a “preference free” model without any ordinal constraints. The SMAA package was then used to conduct to provide the primary outputs for decision making: The rank acceptability index, a vector of central weights, and a confidence factor. The rank acceptability index provides the probability that a treatment is considered best, second best, etc.. under the preference set. The vector of central weights is provided by the preference free model and provides the weights required to have an a priori preference for one treatment over another. The confidence factor is the probability that a given treatment is best when its central weights are used. Lower values of the confidence factor indicate increased uncertainty that a treatment is optimal.

**Sensitivity analysis.** Since SMAA uses absolute event rates, decisions under the same preference set will be sensitive to changes in these rates. We explored the impact of relative changes of +/- 10% in the observed event rates as an approach to assess the sensitivity of the recommended decision.

**Results**

**NMA Results:**

Results from the network meta-analysis were broadly consistent with those in the original publication with some exceptions arising from the use of a joint model. First, since we assume that effectiveness is exchangeable across outcomes, we are able to provide estimates of odds ratios for treatment combinations that did not have data reported (figure x). This exchangeability assumption also means that estimates across outcomes are brought closer together through hierarchical shrinkage. Overall, this effect is modest and while fit statistics do not suggest one model is superior to another (pd model 1 =; pd model 2 =, DIC = xxx vs xxx), we use the joint model to allow for consideration of all treatments across all outcomes.

**SMAA**

Ordinal rankings provided by clinicians are available in Table X. The main differences in ranks provided are due to disagreements on the importance of IVH and oliguria for decision making. The smaa results suggest meaningful differences in recommended treatments and their uncertainty (SM: highest FRA = oral acetaminophen (0.50), CF = 0.66; MCY: highest FRA = IV ibuprofen standard dose (0.31), CF = 0.33) Table X. Central weights for the preference free model suggests that oral acetaminophen is preferred when weights across outcomes are generally equal, while IV ibuprofen standard dose requires heavier weights on intraventricular hemorrhage and oliguria. Both treatments differ from the treatment that performed best on the primary outcome of

**Discussion**

P1 Topic: Summary of results and contrast with original review. No real differences in point estimates using new model but decisions are very different. Contrast with primary/secondary outcome approach

P2/3 Topic: Introduce some literature re: variations in practice and typical response that this a bad thing. Raise the question about whether this may be the result of clinicians trying to synthesize evidence with their values. Expand to parents/families as well. Refer to issue as a sort of double edged sword in that there are questions as to whether the informal assessments people are doing are as reliable as an actual analysis.

**Limitations.**

**-** Model does not account for correlation but inflates variances

- Need to impute missing outcomes. Discuss how the alternative is either to assume the outcome doesn’t matter or that rankings would not change as a result of excluding it. Also look at central weights from preference free model and see if BPD is a big driver.

- Portability. We cannot anticipate every combination of baseline probabilities and clinician preferences.

**Implications for research.** Highlight potential to become a routine addition to NMA. Currently hard to make these accessible to people on the go so we should develop some easy to use tools.

**Implications for practice.** Summarize findings of central weights as well as ordinal (i.e. if these are your beliefs…..)

**Conclusions**

Summarize findings and close with statement that these findings suggest that observed treatment variation may be the result of a rational synthesis of available evidence, local event rates, and outcome preference.