**Use of Multi-criteria Decision Analysis for Treatment Decisions in Pediatrics: A Decision Analysis**

|  |  |
| --- | --- |
| Timothy Disher, PhD(c), RN  Dalhousie University  School of Nursing  Halifax, NS | Souvik Mitra, FRCP, MD  Dalhousie University/IWK Health Centre  Pediatrics  Halifax, NS |
| Chris Cameron, PhD  Cornerstone Research Group  Burlington, ON |  |

**Corresponding Author:**

Timothy Disher, PhD(c), RN

Dalhousie University, School of Nursing

6299, South St

Halifax, NS

Canada

Tim.disher@dal.ca

**Conflicts of Interest:**

**Funding:** No specific funding was received for this project

**Word count:** /3000

**Key Points (75-100 words)**

**Question:**

**Findings:**

**Meaning:**

**Abstract**

**Importance:**

**Objective:**

**Data Sources:**

**Study Selection:**

**Data Extraction and Synthesis:**

**Main outcome and Measures.**

**Results:**

**Conclusions and relevance:**

**Background**

Medical decisions typicaly require patients and health care providers to weigh the pros and cons of multiple outcomes. Within a rational decisoin making framework, this process should take the form of a quantitative weighing of the risk of events multiplied by the desirability or undesirability of the event happening. In even the simple case of two outomes, this process is challenging when decision makers would like to take account of uncertainty in estimates of both outcomes. Difficulty increases as the numbe rof outcomes to consider increase, and the process can quickly become overwhelming. Further complications arise when estimates are generalized to new settings where baseline event rates may differ substantially, requiring new calculations.

A number of alternatives exist within the literature aimed at addressing the issue of optimizing treatment selection in the case of multiple outcomes and incorporating uncertainty. Reviews assessed within the GRADE framework present outcomes in terms of absolute risk reductions alongside their relative reduction counterparts. Qualitative multi-criteria decision aids walk patients through a process of preference elicitation for relevant outcomes. Preference statements typicaly take the form of likert scales, whose results are then summed and qualitatively assessed.

In some cases, these tools are made crudely quantitative through the multiplication of the preference score by the proability of the event and then summing these values. Since the tenants of decision theory suggest that the optimal decision is the one that maximizes expected utility, this process should be provide reasonable decisions. At the most extreme end of the complexity front are methods developed by \*\*XXX\*\* which are embedded within meta-analysis themselves and fully account for uncertainty in ranking, base rates, and relative effects. While this method is attractive from the standpoint of rigour, it can only feasibly be conducted by an expert team, and provides a set of decision weights specific to that team. The issue becomes, for example, when a new clinician in a new setting wishes to adapt these wieghts to their own patients.

Network meta-analyses are a relatively novel synthesis paradigm that allow for all relevant treatments within a connected network to be compared by combining both direct and indirect information. When a treatment comparison contains both direct and indirect copmarison information, network meta-analyses can provide more precise estimates than had the comparison been made based only on the direct comparison. Since NMAs consider all relevant comparators for a given indication, they are attractive for aiding decision making directly. This benefit is further realized by the ability of NMAs to output probabilistic outcomes such as the Surface under the cumulative ranking curve (SUCRA) which places all treatments on the same scale and allows for simple calculatoins for guiding treatment.

Recently, researchers from the London School of Economics developed an interactive decision aid for patients considering statins. The tool is based on the SUCRA output of an NMA, which is weighted by the normalized preferences indicated by patients. While the tool is attractive, it has major theoretical limitations owing to both the insenvitivity of SUCRA to baseline event rates as well as the potential instability of SUCRA itself. First, since SUCRA is based on treatment rankings, treatment weights applied to this outcome will not reflect potentially large difference in the absolute rates of events. Second, SUCRA is typically calculated based on relative effects which themselves can artificially inflate underlying absolute rates. SUCRA further magnifies even small differences in relative risks, potentially overstating the certainty with which one intervention is meaninfully better than another.

The purpose of this project was to develop a simple web based application to facilitate using the results of an NMA or other synthesis product for decicision making accounting for uncertainty. We then compare this approach to decisions based on SUCRA using a recently conducted NMA of treatments for Patent Ductus Arteriosus.

**Method**

**Study Design**

We use a recently published systematic review and meta-analysis of treatments for hemodynamically significant patent ductus arteriosus (PDA) in preterm infants as a case study.

PDA is a common complication in preterm infants (< 37 weeks completed weeks gestation), with potentially serious sequalae including increased risk of serious morbidity and mortality.

Initial pharmacologic treatment consists of non-steroidal anti-inflammatory drugs, most commonly ibuprofen or indomethacin delivered intravenously. Mitra et al conducted a network meta-analysis to compare these interventions as well as acetaminophen in differing doses and routes for the treatment of hemodynamically significant PDA. Elligible trials included randomized clinical trials that enrolled preterm or low birth weight (<2500g) infants.

The review included three primary outcomes and 11 secondary outcomes of which five had sufficient data for quantitative synthesis. Primary outcomes included physical closure or change from hemodynamically signicant to nonsignificant of PDA as determined by echocardiography.

For the purpose of this paper we focus on secondary outcomes that were included in synthesis: death at postmentrual age 36 weeks or before discharge, necrotizing enterocolitis(>= Bell stage 2), bronchopulmonary dysplasia (defined as oxygen use at 36 weeks postmenstrual age), intraventricular hemorrhaghe (any grade by Papile criteria), and oligouria (urine output <1 mL/kg/h). Elligible trials included randomized clinical trials that enrolled preterm or low birth weight (<2500g) infants.

The review included three primary outcomes and 11 secondary outcomes of which five had sufficient data for quantitative synthesis. Primary outcomes included physical closure or change from hemodynamically signicant to nonsignificant of PDA as determined by echocardiography.

For the purpose of this paper we focus on secondary outcomes that were included in synthesis: death at postmentrual age 36 weeks or before discharge, necrotizing enterocolitis(>= Bell stage 2), bronchopulmonary dysplasia (defined as oxygen use at 36 weeks postmenstrual age), intraventricular hemorrhaghe (any grade by Papile criteria), and oligouria (urine output <1 mL/kg/h). A total of 14 treatment nodes were reviewed in 68 studies (N = 4802), although the number of treatments/studies varied considerably across outcomes \*\*(Figure 1 - network graph panel)\*\*. Of interest for this paper are that despite high dose oral ibuprophen performing well across most outcomes, it was among the worst treatments for mortality and oliguria, and performed average for intraventricular hemmorhage. While Mitra et al conclude that this is likely an artifact of the sparsity of the network for these outcomes (e.g. high dose oral ibuprophen provides only one study with a sample size of 30), it should still have implications for decision making. Of further interest is that, despite ranking poorly for PDA closure, wide credible intervals meant placebo or not treatment was not associated with a statistically significant increased odds of death, necrotizing enterocolitis, or IVH.

**Search Strategy and Selection Criteria**

**Study Selection and Data Extraction**

**PDA.** While Mitra et al., stop short of endorsing a treatment recommendation in their reivew, they conclude their discussion by stating that high and standard doses or oral ibuprofen/oral acetaminophen compared to standard regimes.

**Steroids.** Strongest conclusions were to use dexamethasone in early low dose.

**Supplemental oxygen.** Based on editorial and Journal of Pediatrics we extracted all components of the primary outcome separately, ROP requiring treatment, severe NEC, and supplemental oxygen at 36 weeks. Sensitivity analysis conducted based on collapsing components of primary outcome.

**Data sources**

All data was extracted from the base case of the Mitra et al. publication.

These analyses were conducted within a Bayesian framework using netmetaXL, a freely available Excel based software that implements the models described in the NICE TSDs.

Models placed vague priors on treatment effects and the standard deviation of random effects, and properly account of multi-arm trials.

**Statistical Analysis**

We compare these recommendations to those made through two competeting quantitative methods: one based on SUCRA and a second based on absolute probabilities.

**SUCRA method.** The first method is based on the statintool published by \*\*XXX\*\*.

This approach consists of an interactive tool that allows users to rate the importance of outcomes using a slider from 0-100, where zero is unimportant is 100 is vital.

Weights are then normalized to sum to 100 and multiplied against their respective SUCRAs and treatments are ranked on this combined score.

For the SUCRA approach, higher SUCRAs are always better, since they are adjusted to be in the right direction by the original analyst.

**Absolute rate method.** The absolute rate approach is based on standard approaches to MCDA as described by Vandergoert et al.Both SUCRA and absolute rates place events on the same scale however, only absolute rates translate to a rational decision making framework.

For the absolute rates approach importance for "good" outcomes are interpretated as the relative desire to achieve the event so higher absolute values receive higher scores. For "bad" outcomes, ratings are interprated as a statement of preference for avoiding outcomes and is therefore multiplied by the proportion of those not anticipated to to experience the event.

**Sources of uncertainty.** We allow for multiple sources of uncertainty in all models using methods commonly employed in the field of decision modeling. For the SUCRA based tool, we allow for uncertainty in SUCRA ratings themselves (as reported in Mitra et al) as well uncertainty in preference ratings. As these tools are situated within the context of readers of NMAs as opposed to analysts, we assume that they do not have access to original posterior draws of SUCRA.We assume that that SUCRAs can be described by a beta distribution parameterised using the method of moments based on the mean and standard deviation of the Mitra et al paper. For a second source of uncertainty, we allow for users to describe uncertainty in the weights they give to outcomes. This feature can be used to either average multiple sets of ratings together (as in the case of a guideline panel or physician group), or simply to express uncertainty in the exact distribution of weights provided by a single individual. In the latter case, we again use the method of moments to allow users to express a mean weight with a standard deviation and then visualize the implied beta distribution. For the absolute rate method we assume that uncertainty in the relative effectiveness can be represented by a normal distribution on the scale of the linear predictor (e.g. log scale for odds ratios). These estimates are then combined with absolute rates whose uncertainties are defined by a beta distribution. Uncertainty in weights are illustrated using the same methods as the SUCRA approach. For both approaches, all simulations are conducted by simulating 10,000 sets of parameters and calculating expectations of the gain function. Results are displayed as the mean and 95% credible intervals as well as the probability that a treatment is best. \*\*Tornado diagrams for sensitivty to weights\*\*

**Missing data.**

**Results**

**PDA**

**Steroids**

**Oxygen**

**Discussion**

**-** Provide a brief synopsis of key findings with particular emphasis on how the findings add to the body of pertinent knowledge.

- Discuss possible mechanisms and explanations for the findings.

- Compare study results with relevant findings from other published work.

- Discuss the limitations of the study and any methods used to minimize or compensate for those limitations.

- Mention any crucial future research directions.

- Summarize in a straightforward and circumspect manner the clinical implications of the work.

- Limitations

- Ibuprophen high dose oral only ever in three trials max.

- Implication is that despite the large N, you're really making a decision based on those three trials.

- This might be a great case for stronger priors... future work should look at simulations of these situations. Re-analysis outside the scope of this paper though.

- Really should be multivariate

**Limitations**

**Strengths**

**Implications for Research**

**Conclusions**

**Figure Legends**

**Figure 1. Network graphs**. Description

**Figure 2.**

**Figure 3.**

**Figure 4.**

**Figure 5.**

**References**