Ranking method simulation study

Robin Blythe

## Background

This project is an exploratory study to determine whether ranking patients by predicted risk can yield lower costs and better outcomes than using thresholds. Thresholds, while convenient to implement, may be less cost-effective than ranking given that a ranking system should identify true positives first.

A recent consideration with regards to implementing thresholds for risk stratification has been to limit the number of alerts based on what clinicians will tolerate. This appears to be somewhat self-defeating; if only 15 alerts will be tolerated, but the 16th alert is a false negative, this can be severely consequential for patients and hospital safety more generally.

This study contrasts the two approaches - if we randomly sample from a study population, what is the economic impact of applying a ranking approach compared to a threshold approach?

## Study outline

1. Determine a use case.
   1. In this study, we will examine the implementation of a program to identify patients at high risk of deterioration leading to additional costs and worse outcomes.
2. Obtain costs and outcomes for each state: true positive, true negative, false positive, false negative.
3. Simulate a hypothetical, well-calibrated clinical prediction model with a given AUC and underlying prevalence.
4. Simulate an underlying patient population using the minimum required sample size.
5. Identify a threshold based on the entire study population and number needed to treat.
6. Randomly generate new sample populations based on a range of plausible AUC and prevalence values to mimic external validation.
7. Apply the threshold obtained from step 5 and the ranking approach, obtaining costs and outcomes for each sample.
8. Compare approaches across variety of AUCs and prevalence.

## Methods

Let’s assume that, left unassessed, a deteriorating medical inpatient will need to be unexpectedly transferred to the ICU, which incurs significant additional costs. A deterioration detection system can escalate care to the attending doctor; false positives can lead to wasted clinician time, and false negatives can lead to costly ICU admissions.

Considering that many patients will require ICU admission regardless of whether they can be successfully intervened with, we should estimate a conservative treatment effect for assessment.

### Set up experiment

# Hypothetical model specs and event rate  
auc = 0.85  
params = 30  
p0 = 0.01  
  
# Obtain minimum sample size/events  
pmsamp <- pmsampsize::pmsampsize(  
 type = "b",  
 prevalence = p0,   
 cstatistic = auc,  
 parameters = params)  
  
sample\_size <- pmsamp$sample\_size  
min\_events <- ceiling(pmsamp$events)

For a model with 30 parameters, an AUC of 0.85 and a prevalence of 0.01, we need a sample size of 1.3216^{4} with 133 events. We can use these requirements to generate a hypothetical study population.

### Generate simulated dataset

set.seed(888)  
  
# Use model specs and minimum sample size to generate a sample population  
sample\_pop <- predictNMB::get\_sample(  
 auc = auc,   
 n\_samples = sample\_size,   
 prevalence = p0,   
 min\_events = min\_events)  
  
# Obtain predicted probabilities  
fit <- glm(actual ~ x,   
 data = sample\_pop,   
 family = binomial()  
 )  
sample\_pop$predicted <- predict(fit, type = "response")

### Assign costs and outcomes to 2 x 2 table

#### Parameters

* Additional costs due to deterioration ~N(14134, 686) (Curtis et al. 2021)
* Effect of alert on ICU admission ~ $\Beta

#### Ideas:

1. What if we try for a surveillance example too - e.g., inpatient deterioration?
2. Can we add a staff cost for false positives/false negatives?
3. How do we incorporate estimates of staff time in the ranking system?
4. Maybe we quantify how much time clinicians can dedicate to this (e.g. 15 alerts max)
5. Repeat the analysis for a miscalibrated model

Curtis, Kate, Prabhu Sivabalan, David S. Bedford, Julie Considine, Alfa D’Amato, Nada Shepherd, Tiana Elphick, Ramon Z. Shaban, and Margaret Fry. 2021. “Treatments Costs Associated with Inpatient Clinical Deterioration.” *Resuscitation* 166 (September): 49–54. <https://doi.org/10.1016/j.resuscitation.2021.07.022>.