MULTICENTER FLEXIBLE CALIBRATION PLOT WITH BINARY OUTCOMES USING RANDOM EFFECTS

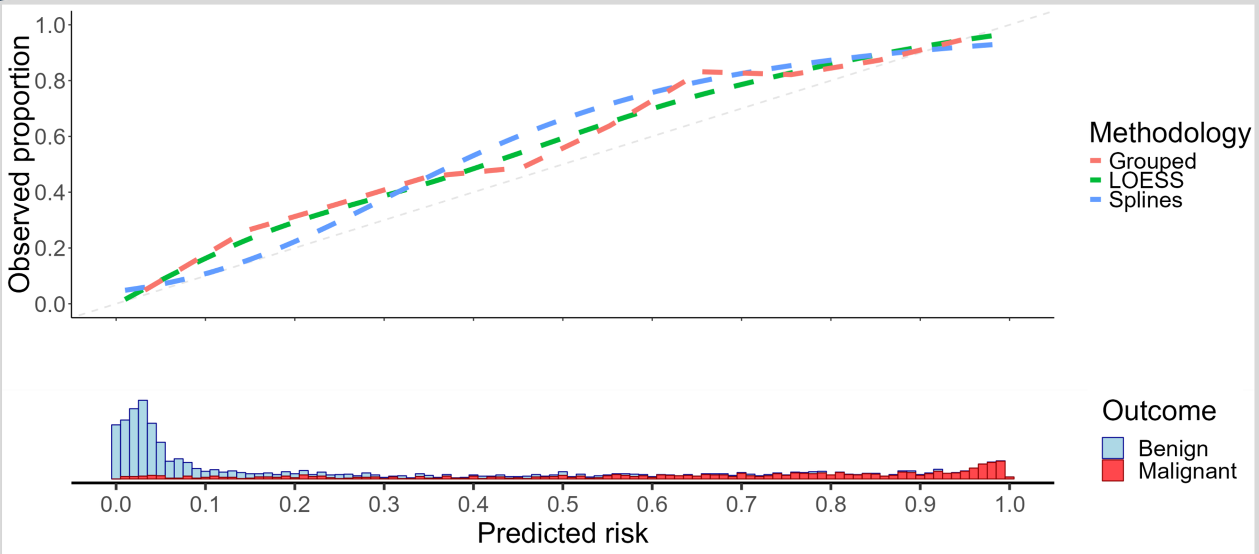
Lasai BARREÑADA[1,2], Laure WYNANTS[1,2], Ben VAN CALSTER[1,2]

[1] Department of Development and Regeneration, KU Leuven, Belgium

[2] Leuven Unit for Health Technology Assessment Research (LUHTAR), KU Leuven, Belgium

Context

- Clinical prediction models predict risk of diagnostic or prognostic events for individuals
- Calibration plots measure the agreement between predicted risks and observed proportion graphically.
- With binary outcomes observed proportion must be estimated from available data.
- It is unclear how to correctly estimate observed proportions and evaluate calibration in clustered data.
- We propose 3 methods to evaluate calibration with calibration curves taking clustering into account using ovarian cancer as a case study.



The calibration plot

- Presents predicted risks in the x-axis and estimated observed proportion in the y-axis.
- Here we present 3 flexible calibration curves
- This calibration plot suggests an underestimation of risks.

Flexible logistic calibration

$$logit(Y) = \alpha + f(LP)$$

Where f is a flexible transformation including:

- LOESS
- Splines (restricted cubic splines)

Clustered quantile meta-analysis

Based on grouping predictions in quantiles

1. Predictions by center (c) are grouped in n quantiles (q_c)
2. The observed proportion is calculated as the event rate in q_c and the predicted risk is the mean predicted risk in q_c
3. Bivariate random effects meta analysis of observed proportion and predicted risk per center and quantile is performed to obtain the pooled result for each quantile.

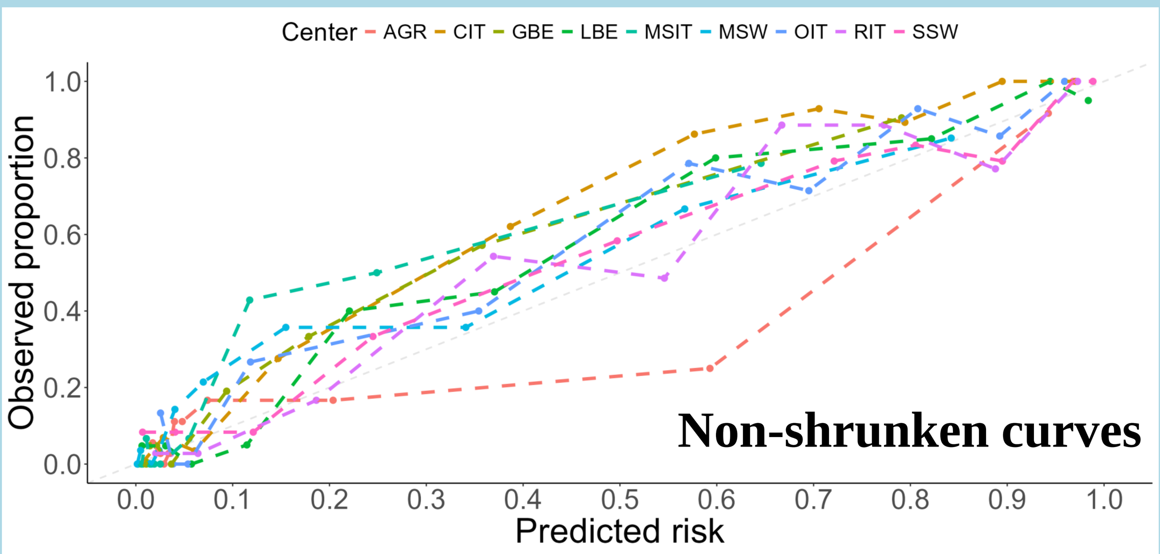


Figure 1. Traditional quantile grouped by center curves

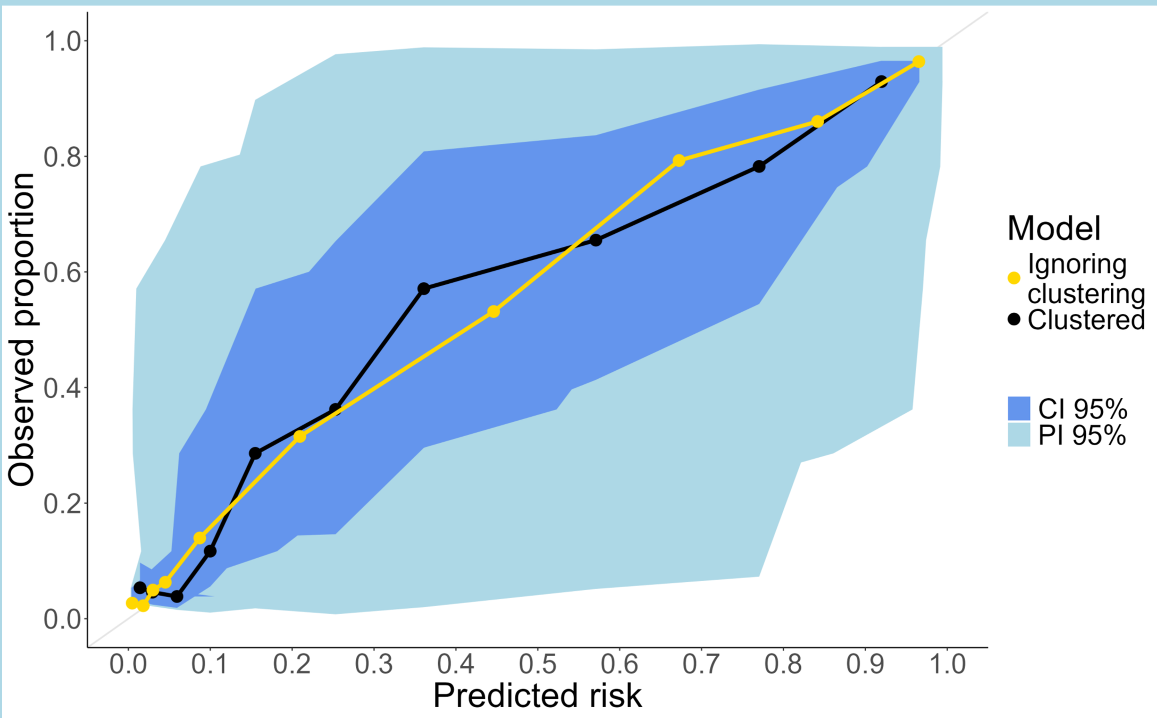


Figure 2. Traditional grouped and clustered quantile meta-analysis calibration plot with 10 groups.

Strengths

- Model agnostic
- Random effects based confidence and prediction intervals
- All clusters present at all groups

Limitations

- Computation time
- Quantiles can contain observations with very disperse risks.
- Curve depends on number of quantiles

Two-stage meta-analysis

Based on pooling center specific flexible curves with random effects meta-analysis

Stage 1: Calculate center specific flexible calibration curves.

Stage 2: Select a grid of 100 fixed predicted risk values. Using the models generated in stage 1 obtain the predicted proportion for the predicted risk value per center. Pool the performance at each value of the grid using random effects meta-analysis.

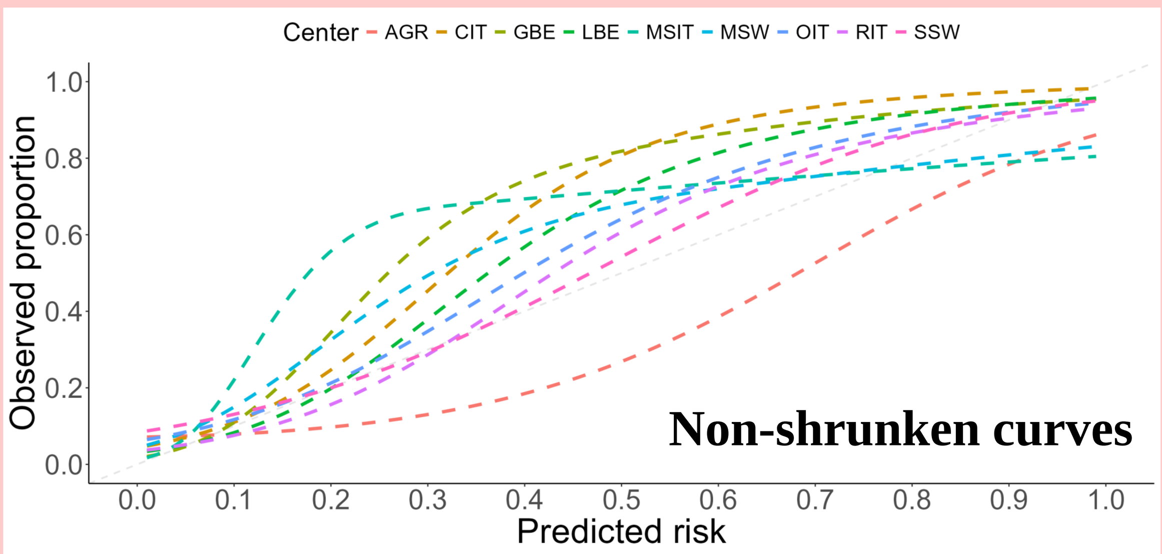


Figure 3. Splines based calibration curves by center

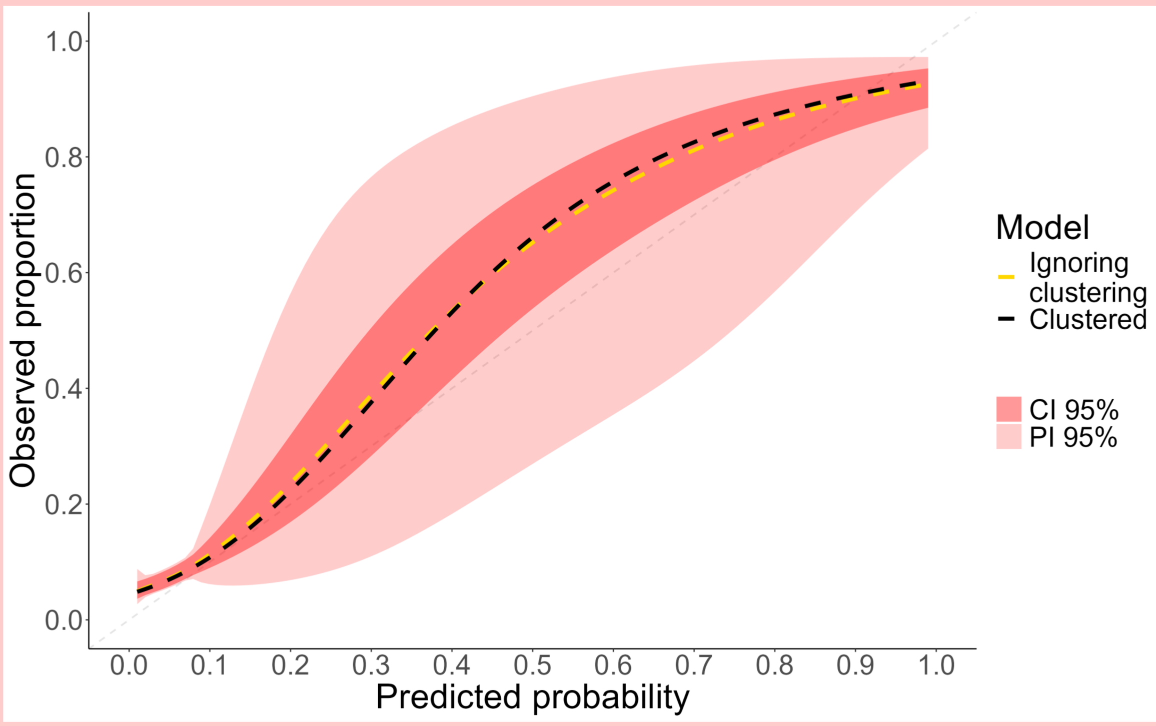


Figure 4. Two stage meta-analysis and splines based calibration curves.

Strengths

- Random effects based confidence and prediction intervals.
- Small centers are accounted for but do not harm the curve.

Limitations

- Computation time
- Two step process
- Curve depends on the smoother used and the smoothing parameters

One stage mixed model

Based on mixed models where the grouping variable is the center. Currently available random intercept and random slope models with splines.

1. Random intercept mixed models estimate an additional coefficient that is the variance of the random effect associated to the center.
2. Random slope models estimate additionally one coefficient per center that is the variance of the random slope.

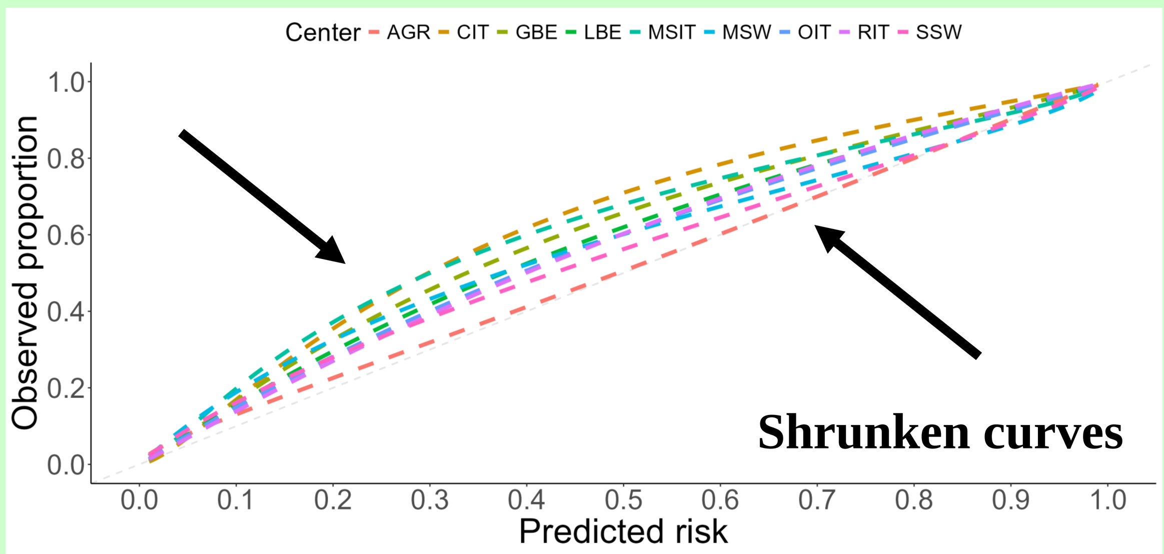


Figure 5. Center specific curves of the random slope model

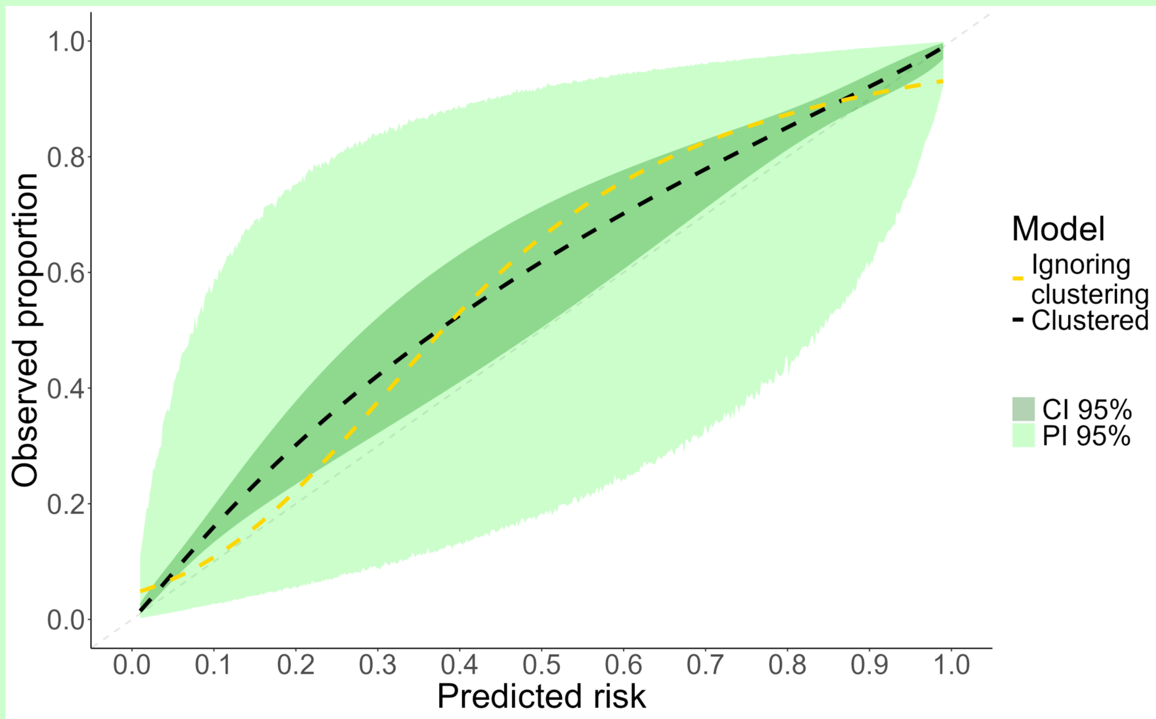


Figure 6. Splines based and random slope mixed model based calibration curves.

Strengths

- Random effects based confidence and prediction intervals.
- One step process.
- Provides shrunken center specific curves.

Limitations

- Computation time
- Requires estimation of one more parameter for random intercept model and one extra parameter per center for random slope model.

@BarrenadaLasai

lasai.barrenadataleb@kuleuven.be