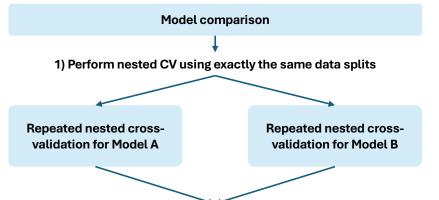


Pool out-of-sample performance estimates (e.g., concordance index) across the external CV folds and across repetitions.

Get 5 x R out-of-sample performance estimates: $[[C_{1,1}, ..., C_{1,5}], ..., [C_{R,1}, ..., C_{R,5}]]$ Compute mean and standard deviation to report.



2) Pool out-of-sample performance estimates across the CV folds and repetitions

$$C^{A} = \left[\left[C_{1,1}^{A}, \dots, C_{1,5}^{A} \right], \dots, \left[C_{R,1}^{A}, \dots C_{R,5}^{A} \right] \right]$$

$$C^{B} = \left[\left[C_{1,1}^{B}, \dots, C_{1,5}^{B} \right], \dots, \left[C_{R,1}^{B}, \dots C_{R,5}^{B} \right] \right]$$

- 3) Perform one-sided paired t-test (C^A , C^B) to test whether mean C^B > mean C^A (or "<" depending on the metrics)
- 4) Report difference in mean estimates and statistical significance (p-value).