Bias in evaluation of discrete surrogate outcomes, due to separation: a penalized likelihood solution

Hannah Ensor*, Christopher Weir

Usher Institute of Population Health Sciences and Informatics, Teviot Place, Edinburgh, EH8 9AG

Surrogate Evaluation

- Surrogate outcomes are biological measures of treatment effect that can be used to predict treatment effect on primary outcomes of interest assessed at a later time point.
- The use of surrogates can reduce the length, size and cost of a clinical trial.
- Crosstabs of binary data with zero cells are known as sparse.
- Bias in estimation can occur in the presence of sparse data.

Methodology

- We focus on the information theoretic approach[1] for surrogate evaluation for a binary surrogate and ordinal true outcome[2].
- The measure R^2_{ht} is calculated via a two stage fixed effects meta-analysis, and lies between 0 and 1.
- R^2_{ht} measures the amount of uncertainty in the true outcome explained by the surrogate at the trial level (the closer to 1 the better the surrogate).

Aims

- Apply a penalized likelihood approach[3] to deal with sparse data within trials. Compare this to a trial removal approach, where trials containing sparse data are discarded.
- Explore overestimation due to overfitting of stage two models. By comparing R^2_{ht} to R^2_{adjC} the standard adjusted R^2 of stage two models (R^2 is consistent with R^2_{ht}).

Simulation Study

- The binary surrogate and ordinal true outcomes are categorised variables that were created as continuous via a bivariate mixed model.
- The table presents results of a simulation study, varying the number and size of trials (R^2_{ht} is set to 0.30).

No. trials	Trial size	Penalized likelihood				Trial Removal approach			
		$Mean\ R^2_{ht}$	$Var\ R^2_{ht}$	$Mean\ R^2_{adjC}$	$Var\ R^2_{adjC}$	$Mean\ R^2_{ht}$	$Var\ R^2_{ht}$	$Mean\ R^2_{adjC}$	$Var\ R^2_{adjC}$
5	10	0.5717	0.0823	0.1386	0.3359	0.6492	0.097	0.0453	0.7192
5	100	0.6122	0.079	0.2244	0.3161	0.6326	0.0794	0.1839	0.447
10	10	0.3597	0.0472	0.176	0.0781	0.6005	0.0848	0.066	0.5203
10	100	0.422	0.0471	0.2569	0.0778	0.4187	0.0521	0.2254	0.0938
20	10	0.2381	0.0197	0.1479	0.0248	0.3957	0.075	0.0255	0.2415
20	100	0.3258	0.0268	0.2464	0.0335	0.308	0.0287	0.2151	0.037
30	10	0.2116	0.0142	0.1529	0.0164	0.3195	0.0497	0.0893	0.0892
30	100	0.3157	0.0173	0.265	0.0199	0.3041	0.0195	0.2455	0.023

Results and Conclusions

- Comparison of R^2_{ht} and R^2_{adjC} *demonstrates the presence of overfitting and overestimation of results particularly for small numbers of trials and weak surrogacy. Weak surrogates are estimated as moderately good in some settings.*
- R^2_{adjC} *removes the associated bias.*
- Comparisons of the penalized likelihood approach and the trial removal approach, using the R^2_{adjC}, *show that the penalised likelihood approach is less biased and more precise.*

❑ *We recommend using a penalized likelihood approach for surrogacy assessment on discrete data.*

❑ *Researchers should also produce the* R^2_{adjC} *to make sure* R^2_{ht} *results are not overly optimistic.*

1. *Alonso, A. & Molenberghs, G. 2006. Surrogate marker evaluation from an information theory perspective. Biometrics, 63, 180-186.*
2. Ensor, H. and Weir C.J. , *Evaluation of surrogacy in the multi-trial setting based on information theory: an extension to ordinal outcomes.* Journal of biopharmaceutical statistics, 2020. **30**(2)
3. Firth, D. (1993) Bias reduction of maximum likelihood estimates. *Biometrika*, **80**, 27–38.

Contact: **hannah.ensor@ed.ac.uk**