



Diagnostic test evaluation: challenges and solutions

Denwood, M.J.¹, Nielsen, S.S.¹, Olsen, A.¹, Comin, A.², Toft, N.³

¹ Department of Veterinary and Animal Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark.

² Department of Epidemiology, Surveillance and Risk Assessment, Swedish Veterinary Agency, Uppsala, Sweden. ³ Toft Analytics ApS, Værløse, Denmark.

Motivation

- Diagnostic test evaluation using latent class models is a recurring theme in many veterinary and human medical applications
- Work is ongoing to apply latent class models in the context of neglected tropical diseases (NTDs), which are typically characterised by complex life cycles of pathogens and relatively poor sensitivity of existing reference tests
- The applicable methods are well understood by the statistical community, but challenges remain when applying them in practice

Determining the Latent Class

Challenge: interpretation of sensitivity and specificity is always relative to an implicit latent class, which is defined by the combination of tests and populations used to fit the model. It is not always clear from published studies when this latent class does or does not match the desired target condition, which severely limits the practical utility of the estimates provided.

Solution: a directed acyclic graph (DAG; Fig. 1) is an excellent way to visualise the biological processes relating the tests used, from which the latent class can be more easily deduced. The DAG should start with the desired target condition at the top, branch through biological processes as intermediate nodes, and end with the observed test results at the bottom. This is particularly relevant for NTDs, where the biology tends to be complex.

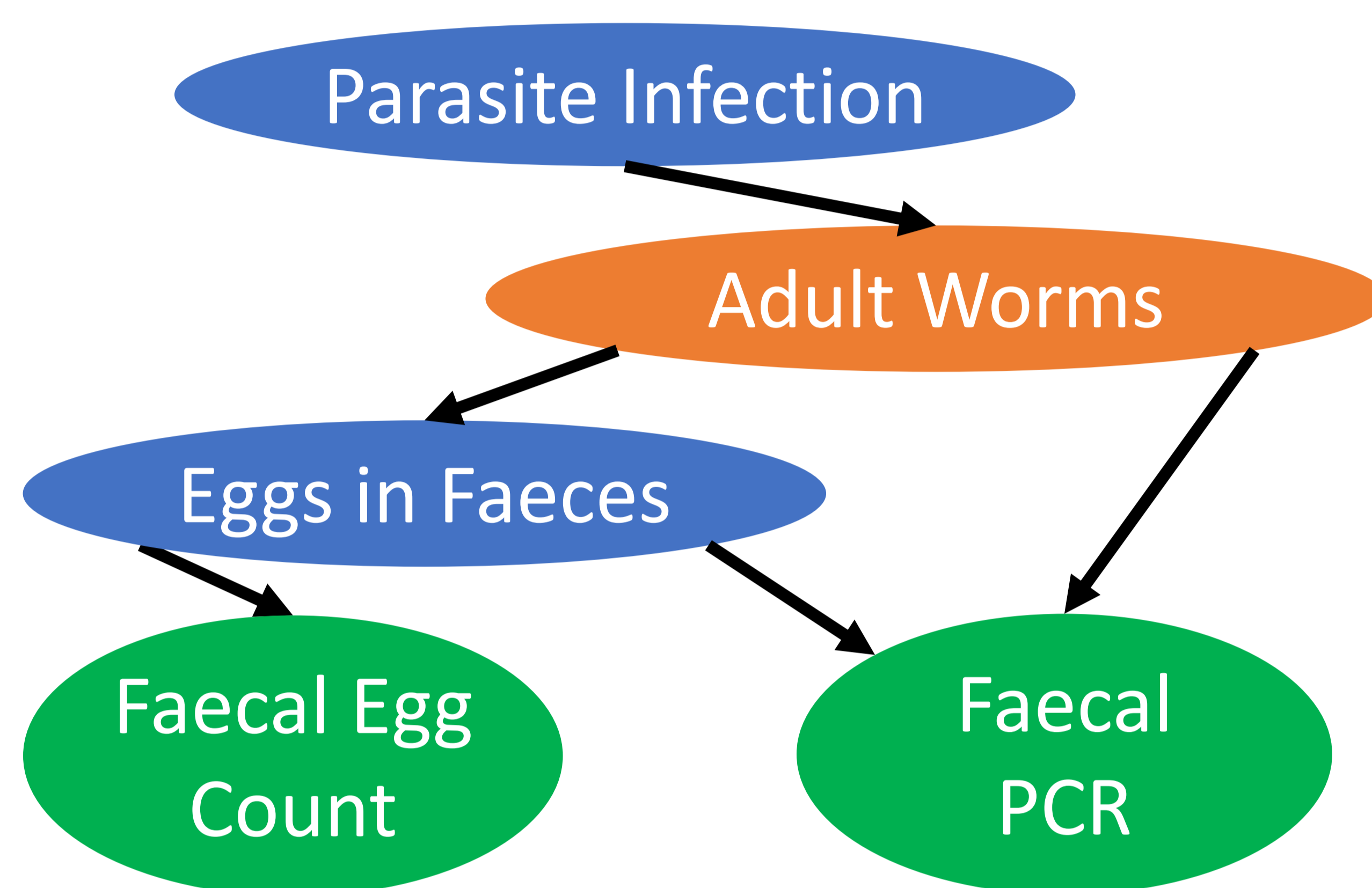


Fig. 1: Illustrative example of a directed acyclic graph (DAG) for two tests (green) and a number of underlying biological processes (blue), indicating that the implicit latent class is presence of adult worms in the gut (orange)

Model Fit Assessment

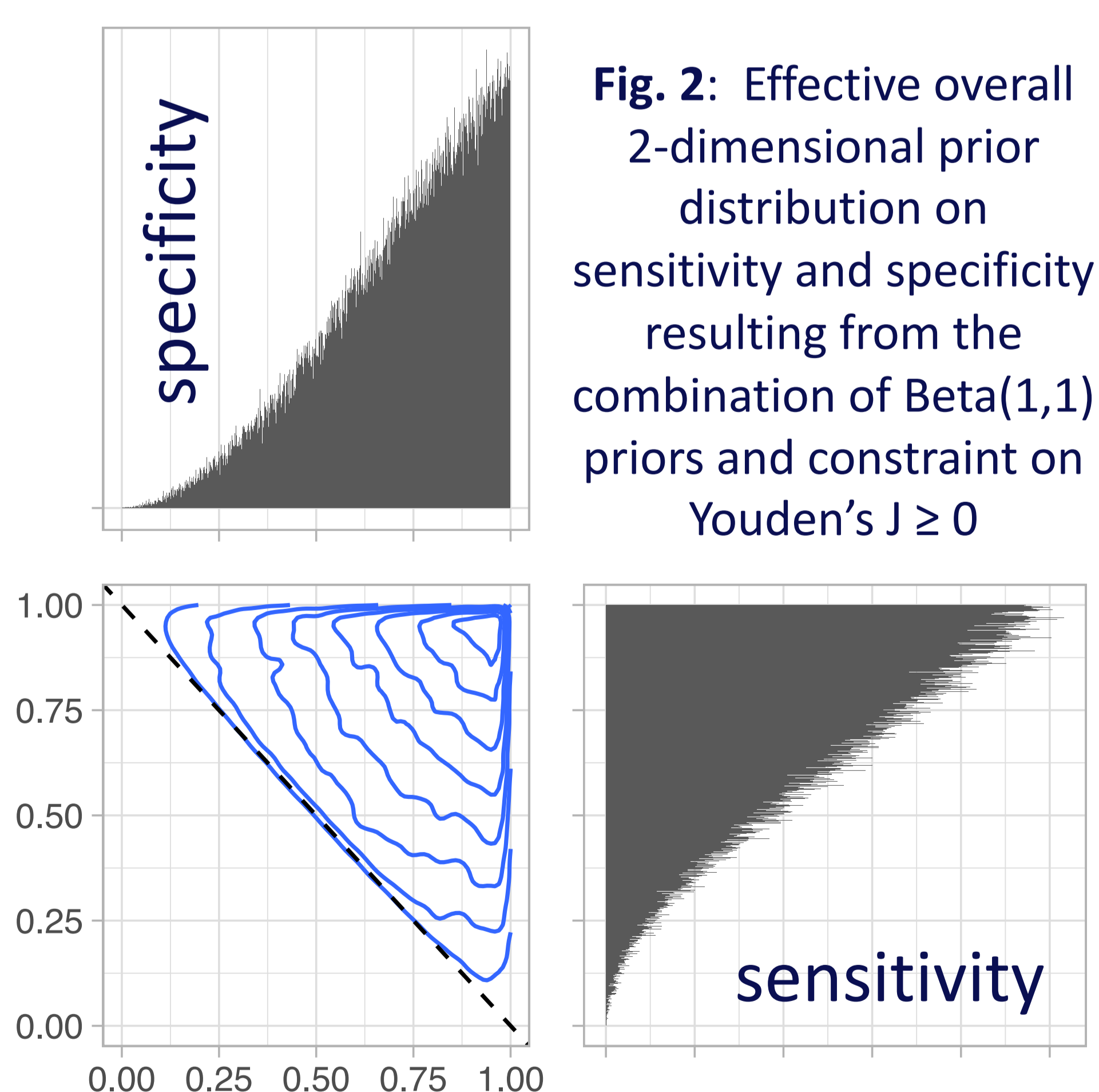
Challenge: we must ensure that important assumptions are not violated, including constant test performance across populations.

Solution: some methods of checking empirical fit and model assumptions are provided by `runjags::template_huiwalter`¹

Identifiability and Constraints

Challenge: parameterisation of classic Hui-Walter models should be constrained to ensure (1) Youden's J (sensitivity+specificity-1) is above zero for each test so that label switching does not occur, and (2) conditional dependence terms do not result in invalid probabilities. Attempting to fit too many conditional dependence terms results in an unidentifiable model.

Solution: MCMC and most HMC algorithms (i.e. JAGS and Stan) can include an explicit accept/reject step to impose necessary constraints – the model formulations produced by `runjags::template_huiwalter`¹ include the necessary terms. Combining Beta(1,1) priors with constraints on Youden's J results in effective priors that resemble Beta(2,1) distributions (Fig 2). Choice of conditional dependence terms should reflect known branching within the DAG and should be restricted to only those that are necessary. Identifiability of the model should be verified before obtaining a model fit even if MCMC sampling appears to produce reasonable results. Two-test models should never incorporate conditional dependence.



Conclusions

- Drawing a DAG is an extremely useful way of visualizing and demonstrating the latent class and potential conditional dependencies
- More attention should be given to demonstrating identifiability and empirical fit of models (particularly those fit using MCMC)
- Priors for sensitivity and specificity should be considered in combination and given the relevant model constraints

¹ <https://CRAN.R-project.org/package=runjags>