Benefits of Bayesian Priors as Regularizers in Challenging Meta-analysis Settings

Emma Mackay
Scientific Advisor - Statistics
Inka Health

JSM 2024 – Portland, OR

Disclosures

- Advisor at Inka Health, an HEOR analytics consulting company
- No conflicts of interest related to this presentation
- Opinions expressed are my own

Motivation

- The drive for precision medicine is resulting in drug development targeting increasingly narrow patient populations
  - Leading to challenges with limited studies meeting screening criteria for meta-analysis and often more limited sample sizes within-study
  - Increasing difficulty for some types of meta-analysis (e.g. sparse networks in NMA and difficulties estimating complex random effects models)
- Presentation goal: to outline how use of weakly informative priors under a Bayesian meta-analysis framework can be beneficial for mitigating issues of overfitting and weak identifiability in challenging meta-analysis settings
- Will focus on two meta-analysis examples as illustration

Surrogate Endpoint Validation

- Key question: will therapies that improve surrogate endpoints like progression-free survival (PFS) yield improvements in more crucial endpoints?
  - E.g. will improvements in PFS translate into later improvements in overall survival (OS)?
  - An association between surrogate and clinical endpoint is not sufficient as treatment effects may not be mediated through the surrogate[1]

- More rigorous criteria for establishing the validity of a surrogate endpoint have been outlined (for example, see Buyse et al.[2] and guidance from the NICE DSU[3])

[1] Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled?. Annals of internal medicine. 1996.

[2] Buyse M, Molenberghs G, Burzykowski T, Renard D, Geys H. The validation of surrogate endpoints in meta-analyses of randomized experiments. Biostatistics. 2000.

[3] Bujkiewicz S, Achana F, Papanikos T, Riley RD, Abrams KR. NICE DSU Technical Support Document 20: Multivariate meta-analysis of summary data for combining treatment effects on correlated outcomes and evaluating surrogate endpoints. 2019.

Meta-Analytic Surrogate Endpoint Validation

- Basic idea: do RCTs that find treatment benefits on surrogate endpoints also tend to find benefits on more crucial endpoints?
- Meta-analytic approaches focus on estimating the cross-trial association between treatment effects on the surrogate and true endpoints via (weighted) linear regression methods (often under a bivariate normality assumption)
- Since focus is on cross-trial variation (rather than pooling treatment effects across multiple study populations) sample sizes tend to be very small

- Will focus on the Daniel & Hughes meta-analytic specification[4]:

$$\begin{pmatrix} Y_{1i} \\ Y_{2i} \end{pmatrix} \sim N\left(\begin{bmatrix} \delta_{1i} \\ \delta_{2i} \end{bmatrix}, \begin{bmatrix} \sigma_{1i}^2 & \rho_{wi}^{12}\sigma_{1i}\sigma_{2i} \\ \rho_{wi}^{12}\sigma_{1i}\sigma_{2i} & \sigma_{2i}^2 \end{bmatrix} \right) \quad (1)$$

$$\delta_{2i}|\delta_{1i} \sim N(\lambda_0 + \lambda_1\delta_{1i}, \psi^2) \quad (2)$$

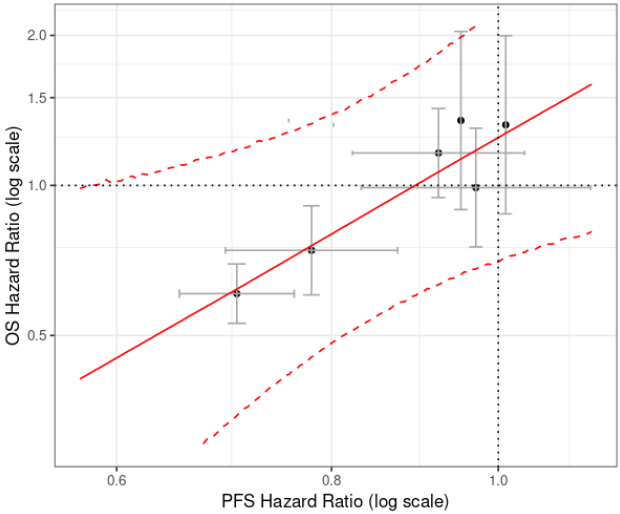
where the effect estimate on the surrogate, Y_{1i}, and effect estimate on the true endpoint, Y_{2i}, in each study i are realizations from a bivariate normal with surrogate and clinical endpoint effect parameters δ_{1i} and δ_{2i}, respectively, within-study correlation ρ_{wi}^{12}, and standard errors σ_{1i} and σ_{2i}.

- The main component of interest in the model is the study-level linear regression model in (2), with the λ_1 coefficient capturing the strength of the surrogacy relationship (between-study association).

[4] Daniels MJ, Hughes MD. Meta-analysis for the evaluation of potential surrogate markers. Statistics in medicine. 1997.

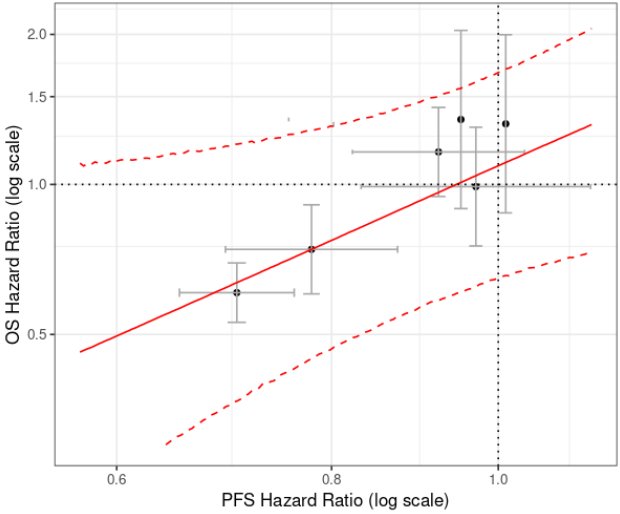
Toy Example with Diffuse Priors

- Daniel & Hughes model with diffuse priors applied to 6 studies reporting hazard ratios (HR) for PFS (surrogate) and OS (true endpoint)
- Perhaps concerning that
  1. intercept suggests that, on average, studies with no benefit in terms of PFS would be expected to have an OS HR > 1.2
  2. the slope estimate is largely driven be two studies
  3. and yet a new study with a PFS HR slightly below 0.6 would be predicted to have an OS benefit based on a the 95% prediction interval



Toy Example Revisited?

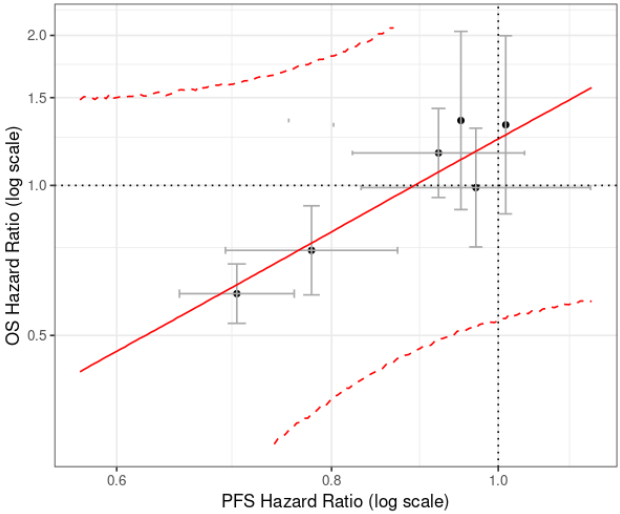
- Generally, we would expect the line of best fit to pass through the origin (log HR of 0 for both PFS and OS)[5]
- Can perhaps mitigate overfitting when degrees of freedom are limited by choosing priors which shrink the intercept towards zero
- Righthand figure imposes a prior $\lambda_0 \sim N(0, 0.1)$



[5]Note that there are, however, some exceptions

Toy Example Revisited?

- May also want to consider using a prior on ψ^2 (the parameter for the variance around the regression line) that avoids values close to zero where the number of studies is severely limited
- Righthand figure imposes a prior $\psi \sim \textit{Gamma}(2, 1)$ (right-skewed with mode above zero)



- The performance of diagnostic tests across studies can be suspectible to not only differences in patient populations but also differences in thresholds for test positivity
- Focus will be on a bivariate random effects meta-analysis approach
- Similar challenges arise in modelling cross-study heterogeneity if we identify few studies evaluating the diagnostic test of interest in our SLR (i.e. overparameterization)

- We follow simplified version of the parameterization used by Verde[6,7]:

$$
\begin{aligned}
tp_i &\sim \text{Binom}(n_{i1}, TPR_i) \\
fp_i &\sim \text{Binom}(n_{i2}, FPR_i) \\
D_i &= \text{logit}(TPR_i) - \text{logit}(FPR_i) \\
S_i &= \text{logit}(TPR_i) + \text{logit}(FPR_i)
\end{aligned}
$$

$$
\begin{pmatrix} D_i \\ S_i \end{pmatrix} \sim N\left(\begin{bmatrix} \mu_D \\ \mu_S \end{bmatrix}, \begin{bmatrix} \sigma_D^2 & \rho\sigma_D\sigma_S \\ \rho\sigma_D\sigma_S & \sigma_S^2 \end{bmatrix} \right)
$$

,

	With Disease	Without Disease
Test +	tp_i	fp_i
Test −	fn_i	tn_i
Sum	n_{i1}	n_{i2}

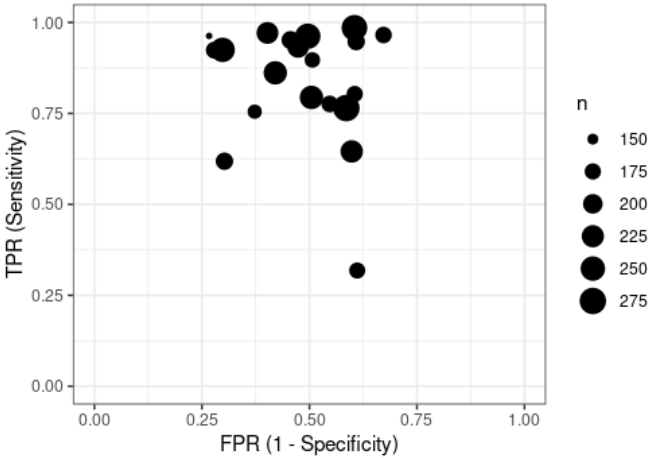
where TPR_i is the true positive rate and FPR_i the false positive rate for study i (with empirical values tp_i/n_{i1} and fp_i/n_{i2}, respectively)

[7] Verde PE. bamdit: An R package for Bayesian meta-analysis of diagnostic test data. Journal of Statistical Software. 2018.

[7] Verde PE. Meta-analysis of diagnostic test data: a bivariate Bayesian modeling approach. Statistics in Medicine. 2010.

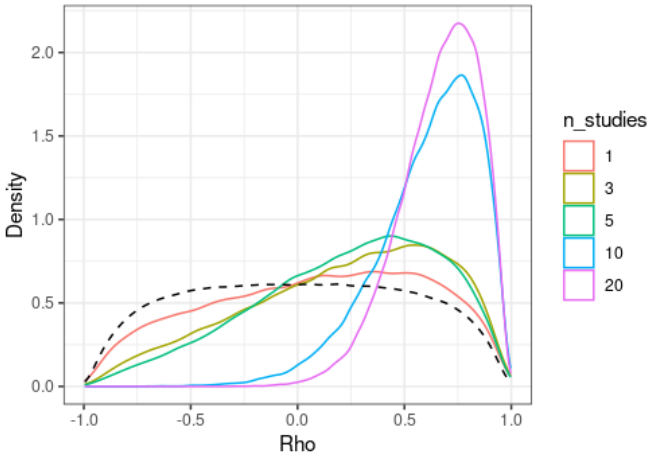
Another Toy Example

- We simulate data under the bivariate random effects model with correlation coefficient $\rho = 0.7$
- We use weakly informative priors suggested by Verde, including $z \sim N(0, 1.7)$ with $z = logit(\frac{\rho+1}{2})$ (see next slide)



Another Toy Example

- Consider the cross-study correlation parameter between D (diagnostic odds ratio) and S, ρ, for example
- Weakly informative priors allow for estimation even when model is overparameterized (e.g. $n_{studies} = 1$)
- Posterior reflects agnostic beliefs expressed in prior (black dashed line) when there is minimal evidence to update our beliefs



Food for Thought

- Use of weakly informative priors can potentially be used to:
  - Mitigate overfitting in in data-scarce meta-analysis settings
  - Allow for interpretable results even in settings of overparameterization or weak identifiability
- Priors can be specified in advance of data collection (conducting of SLR) or weakly informative priors could be imposed as a sensitivity analysis
- Unfortunately it may not always be trivial to formulate reasonable default priors. Regardless, we should at least be mindful of whether our analysis plan is equipped to handle small sample sizes and/or few studies
- Additionally, model stability / MCMC convergence can be a real challenge with extremely small sample sizes–not always trivial to implement!

Thank You!

Contact: Emma Mackay,
emma@inka.health