Bayesian Borrowing Approaches to Address the Challenges of Evaluating Efficacy/Effectiveness in Rare Indications: *Applications to Basket Trials and Pediatric Studies*

Emma Mackay
Scientific Advisor - Statistics
Inka Health

SSC 2024 St. John's, NL

Disclosures

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

Background (1 / 2)

- Rare disease settings provide a particularly challenging setting for evaluating the efficacy of new drugs:
  - Can be difficult to recruit enough patients to run a conventional well-powered randomized controlled trial (RCT)
- Growing use of unconventional methods, e.g.:
  - Borrowing from historical trials to augment small concurrent control arms
  - Fully-externally controlled trials
  - Newer trial designs like basket trials (recruit patients with multiple disease subtypes who share a common druggable target–e.g. cancer mutation)

Background (2 / 2)

- Small sample size challenges also arise in many pediatric trials–can we borrow information from similar trials in adult populations?
- Increased focus on 'precision medicine' in drug development as patient population for new treatments becomes increasingly narrowly defined (e.g. patients with a specific cancerous genetic mutation)
- Growing receptiveness to the use of Bayesian borrowing methods and synthetic or hybrid control arms where conventional trials are impractical/infeasible[1,2,3]

[1] US FDA. Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials. Guidance for Industry and FDA Staff. 2010.

[2] US FDA. Interacting with the FDA on Complex Innovative Trial Designs for Drugs and Biological Products: Guidance for Industry. 2020.

[3] US FDA. Adaptive Designs for Clinical Trials of Drugs and Biologics Guidance for Industry. 2019.

Challenge

- Rationale of conducting an RCT:
  - Randomize patients to experimental or control treatment so that only treatment received differs systematically between treatment arms $\rightarrow$ Allows us to infer causal effect of treatment assignment on outcomes
  - Can design trial to achieve type-I and type-II error operating characteristics
- Where it is difficult or infeasible to design a well-powered RCT, we can borrow information from data sources external to the trial, however this:
  - Introduces risk of bias as external data sources are not subject to randomization (differences in patient populations other than treatment received risk confounding treatment effect estimates)
  - Makes it difficult to achieve target type-I and type-II error operating characteristics

How to Approach Borrowing from External Data Sources (1 / 2)

- Care needed in identification of external data source (e.g. historical control arm)
  - Similar patient population (e.g. similar eligibility criteria for historical control, and similar baseline characteristics)
- Appropriate application of methods for adjusting for observed differences in potential confounders
  - Could control for confounds using parametric model (outcome regression)
  - Could construct synthetic control arm with similar baseline characteristics using matching or inverse probability of treatment weighting (IPTW) methods

How to Approach Borrowing from External Data Sources (2 / 2)

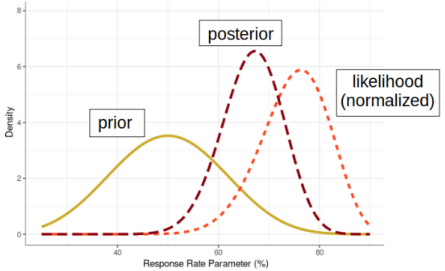
- But what about residual heterogeneity across populations/data sources?
  - Can down-weight the information contribution of the external data when there is evidence of meaningful differences in (post-adjustment) outcomes between data sources $\rightarrow$ "dynamic borrowing'
    - E.g. borrow less information from an external control to supplement a small concurrent control arm in an RCT
    - Or when estimating, say, the disease control rate for the control arm for a fully-externally controlled trial, heterogeneity in outcomes across external data sources should be reflected in the precision of estimate of the disease control rate
  - Focus of this presentation

Why Use a Bayesian Approach for Rare Diseases?

- Provides a principled framework for incorporating external information:
  - Start with our prior (which can be informed by external data)
  - Update our beliefs after observing new data
  - Conducive to sequential 'Bayesian updating'
  - Posterior inference allows us to quantify the amount of evidence in favour of a conclusion and allows for more nuanced decision rules
  - See Mackay & Springford (2023) for additional discussion[4]

$$p(\theta \mid \text{Data}) = \frac{p(\text{Data} \mid \theta) \cdot p(\theta)}{p(\text{Data})}$$

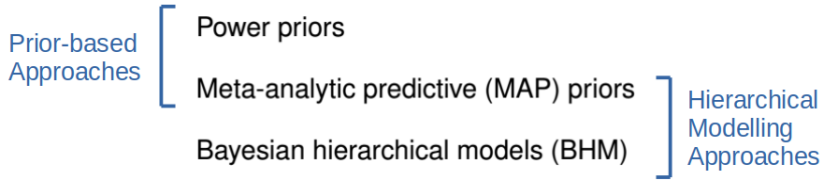
posterior; likelihood (contribution of new data); prior; normalizing constant



[4] Mackay EK, Springford A. Evaluating treatments in rare indications warrants a Bayesian approach. Frontiers in Pharmacology. 2023.

Methods Roadmap

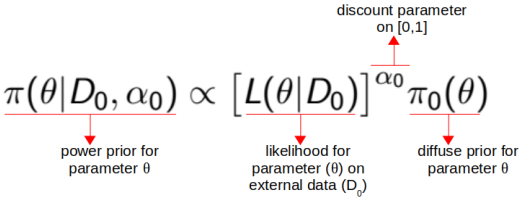
- Will cover three approaches to Bayesian borrowing:



- Typically external data is only available for control treatments and this data (can use external data form a prior for the control arm parameter or could use it as a stand-in for a non-existent control arm)
- Without loss of generality, examples will focus on binary response endpoints (will use external data to inform estimates of the control treatment response rate)

Power Priors

- Power priors[5]
  - Down-weight the external data by means of a discount parameter, $\alpha_0 \in [0, 1]$



  - As $\alpha_0 \to 0$ we ignore the external data (no pooling)
  - As $\alpha_0 \to 1$ we give it full weight (full pooling)
  - How to choose α_0?: 'tipping point' approach[6], target effective sample size for borrowing[7], dynamic borrowing based on consistency between data sources[8]

[5] Ibrahim JG, Chen MH. Power prior distributions for regression models. Stat. Sci. 2000.

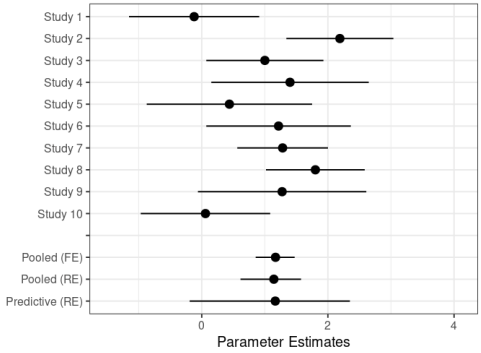
[6] Best N, ... Assessing efficacy in important subgroups in confirmatory trials: An example using Bayesian dynamic borrowing. Pharm. Stat. 2021.

[7] Richeldi L, ..., Maher TM. Trial of a preferential phosphodiesterase 4B inhibitor for idiopathic pulmonary fibrosis. NEJM. 2022.

[8] Ibrahim JG, Chen MH, Gwon Y, Chen F. The power prior: Theory and applications. Statistics Med. 2015.

Meta-analytic Predictive (MAP) Priors (1 / 2)

- Meta-analysis approach to construct an informative prior (e.g. for the average response under a SoC treatment)[9]
- Since response may vary across trial populations, we want our prior to incorporate both within-trial and between-trial uncertainty
- Idea is to conduct a random-effects meta-analysis and use the posterior predictive distribution (predicted SoC response in a new trial) as our prior



[9] Neuenschwander B, Capkun-Niggli G, Branson M, Spiegelhalter DJ. Summarizing historical information on controls in clinical trials. Clinical Trials. 2010.

Meta-analytic Predictive (MAP) Priors (2 / 2)

Basic Setup

- Have control arm response data $D_h = (Y_h, n_h)$ for $h = 1, ..., H$ historical trials
- Assume that $\theta^*, \theta_1, ...\theta_H \sim N(\gamma, \sigma^2)$ where θ^* is a the log-OR in a hypothetical new trial
- Can compute the MAP prior for parameter θ in our concurrent control arm as the posterior predictive distribution $[\theta^*|D_1, ..., D_H]$

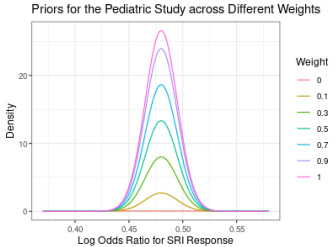
- Low cross-trial heterogeneity $\rightarrow$ greater effective sample size borrowed
- Can be made more robust to prior-data conflict (Robust MAP) by using a weighted mixture between the MAP prior (f_{MAP}) and a vague prior (f_V)[10]
  - $w \cdot f_{MAP}(\theta) + (1 - w) f_V(\theta)$

[10] Schmidli H, Gsteiger S, Roychoudhury S, O'Hagan A, Spiegelhalter D, Neuenschwander B. Robust meta-analytic-predictive priors in clinical trials with historical control information. Biometrics. 2014.

- Post-marketing pediatric study required by FDA to evaluate efficacy of belimumab in pediatric SLE patients for SLE responder index (SRI) endpoint
- Analysis to supplement limited pediatric trial population by borrowing from adult trials via robust MAP prior
- Informative prior for pediatric log odds ratio of SRI response (δ) was constructed using a robust MAP approach following a meta-analysis of two adult studies:

$$\delta \sim (1 - w) \cdot N(0, 8.27^2) + w \cdot N(0.48, 0.015^2)$$

- Assess weight w on MAP component required to reach efficacy tipping point
- Concluded that amount of borrowing from adult trials to reach tipping point was acceptable, leading to approval



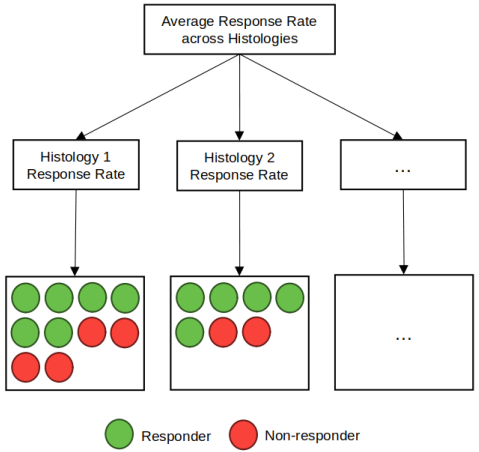
Source data from: `https://www.fda.gov/media/127912/download`

Basket Trials

- Basket trials recruit patients with multiple disease subtypes (e.g. lung cancer, breast cancer) as long as they are positive for the mutation/biomarker that the experimental drug targets
  - These trials are usual single-arm (lack a control arm)
  - Typically done to increase sample sizes where the targeted mutation/biomarker is very rare
- To pool the data or not to pool?
  - Complete pooling ignores potential heterogeneity in response across tumour types $\rightarrow$ results may not generalize
  - No-pooling $\rightarrow$ back to problem of small sample sizes

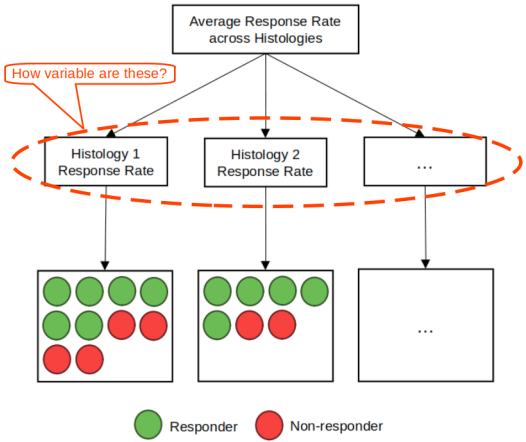
- Bayesian hiearchical models (BHM) allow for partial pooling–a middle-ground between the extremes of complete pooling and no pooling
- Allows response rates to differ across histologies but assumes they are related (“exchangeability assumption”)
- Amount of partial pooling (or “borrowing”) across histologies depends on degree of heterogeneity in responses across histologies

Bayesian Hierarchical Model



- Heterogeneity parameter is estimated based on the trial data
- High heterogeneity $\rightarrow$ little borrowing
- Low heterogeneity $\rightarrow$ more borrowing

Bayesian Hierarchical Model



Example from Murphy et al. (2021)[11]

- Analysis of a basket trial for larotrectinib in NTRK-fusion-positive solid tumours
- Model for histologies $k = 1, ..., K$:

$$r_k \sim \text{Binom}(n_k, p_k)$$
$$\text{logit}(p_k) = \theta_k$$
$$\theta_k \sim N(\mu, \sigma^2)$$

$$\text{priors: } \mu \sim N(-0.8473, 10)$$
$$\sigma \sim \text{Unif}(0, 5)$$

- Partial pooling yields histology-specific response rates that are shrunken towards the average–particularly in the case for tumour types with very few patients

Probabilities of Response for Each Histology

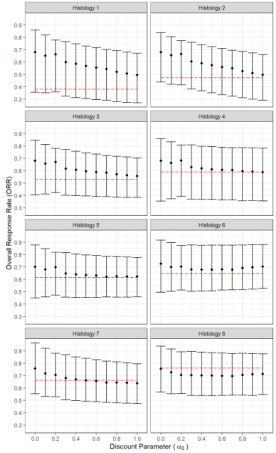
Histology	Observed Response	Estimated Response under BHM Mean (95% CrI)
Soft-tissue sarcoma	10/11 (90.9%)	88.1% (66.0% – 99.1%)
Salivary gland	10/12 (83.3%)	81.8% (58.0% – 96.8%)
IFS	7/7 (100%)	93.3% (70.5% – 100%)
Thyroid	5/5 (100%)	91.6% (63.0% – 100%)
Lung	3/4 (75.0%)	72.6% (30.4% – 97.8%)
Melanoma	2/4 (50.0%)	52.5% (12.4% – 89.4%)
Colon	1/4 (25.0%)	32.0% (2.6% – 75.5%)
GIST	3/3 (100%)	88.3% (49.3% – 100%)
Cholangiocarcinoma	0/2 (0%)	21.0% (0.0% – 75.7%)
Appendix	0/1 (0%)	30.0% (0.1% – 89.7%)
Breast	0/1 (0%)	30.0% (0.1% – 90.1%)
Pancreas	0/1 (0%)	29.8% (0.1% – 89.7%)

[11] Murphy P, Claxton L, Hodgson R, Glynn D, Beresford L, Walton M, Llewellyn A, Palmer S, Dias S. Exploring heterogeneity in histology-independent technologies and the implications for cost-effectiveness. Medical Decision Making. 2021.

Combined Approach Example from Mackay et al. (2023)[12]

- Uses a combined BHM and power prior to facilitate:
  1. Partial pooling of information across histologies under an exchangeability assumption, and
  2. Partial borrowing from adult basket trial data to supplement pediatric trial data using a power prior
- We demonstrate the approach using simulated data under a scenario where borrowing from the adult data is clinically appropriate
- Righthand figure shows how overall response rate (ORR) estimates change with increased borrowing weight on the adult data

Figure 1. Pediatric ORR Estimates (Posterior Median and 95% CrIs) for Various Borrowing Weights (Red Line Indicates True ORRs)



[12] Mackay E, Springford A, Heeg B, Arora P, Thorlund K. Combating Sample Scarcity: A Novel Bayesian Approach to Pediatric Basket Trials [Abstract]. Value in Health. 2023. https://doi.org/10.1016/j.jval.2023.09.2141

Concluding Remarks

- Bayesian borrowing approaches present a structured way to leverage all available external data sources when faced with severe data limitations in the evaluation of rare diseases
- Growing receptiveness to their use in complex and innovative clinical trial designs
- These methods can allow for flexible incorporation of disparate data sources under different structural modelling assumptions (e.g. aggregate-level and individual patient data)
- However extreme care needs to be taken to assess (i) suitability of data sources, (ii) structural modelling assumptions, and (iii) and sensitivity of results to key inputs (priors, etc.)

Thank You!

Contact: Emma Mackay,
emma@inka.health