Cytel

A Novel Information Borrowing Approach for Evaluating Response in Pediatric Basket Trials with Limited Sample Sizes

Bayesian Biostatistics Conference 2023, Utrecht, Netherlands

Presenter: Emma Mackay
Coauthors: Aaron Springford, Paul Arora, Bart Heeg and Kristian Thorlund
Date: October 26, 2023

Background

Basket trials → patients with multiple disease subtypes & common targetable mutation/biomarker

- Challenge: differential response by disease subtype?
  - To pool or not to pool?

Basket trials are seeing increased use

- Assess safety and efficacy of novel treatments targeting very rare cancer mutations [1]

Precedent for information borrowing methods in rare indications

- Bayesian hierarchical modelling (BHM) proposed to account for heterogeneity in outcomes across disease subtypes [2,3]
  - Allows for partial pooling of information across disease subtypes (e.g. tumour histologies) based on degree of heterogeneity in outcomes
- Recent application supplementing limited sample sizes in pediatric trials using data from adults [4]
- A key idea is to borrow when data are “compatible” and avoid substantial borrowing when data are “incompatible” [5]

Limited sample sizes in pediatric single-arm basket trials

- Benefit to (1) partial pooling of information across histologies, and (2) information borrowing from adult basket trials

What Should a Desirable Approach Include?

- Reduce pooling across histologies when response rates are heterogeneous
- Reduce borrowing from adults when response rates are different between adult and pediatric populations
- Allow for detailed sensitivity analysis*
- Allow for different structural assumptions based on clinical input on anticipated heterogeneity in response across histology and age (adult vs. pediatric):

Is the degree of heterogeneity in response across histologies likely to be similar between pediatric and adult populations?	Is response anticipated to be generally higher/lower in pediatric populations?	Are histologies that respond well to treatment in pediatric patients expected to respond well in adult patients too?	Based on clinical knowledge, are there some histologies that should be analyzed completely separately?

*e.g. tipping point analysis [4,6], and prior sensitivity analyses

Objectives

- To outline a Bayesian hierarchical modelling (BHM) framework that can allow for information borrowing:

  1) across histologies, and;

  2) from adult populations

  ...that can improve precision of efficacy estimates in pediatric basket trials.

- Demonstrate multiple approaches and trade-offs via simulation

Proposed Model Setup

- We focus on a binary response endpoint with outcomes $y_i \in \{0, 1\}$ for $i = 1, \ldots, n_P + n_A$ pediatric and adult participants, respectively, who have one of $h \in \{1, \ldots, H\}$ histologies / tumour types
- We assume that response is determined according to the following Bernoulli model:

$$y_i \sim \text{Bernoulli}(p_i)$$

$$\text{logit}(p_i) = \mu + \gamma_{h(i)} + \left(\eta_{h(i)} + \delta\right) \cdot 1\{c_i = 1\}$$

- Where c_i is a 0/1 indicator for membership in the adult cohort. $\gamma_{h(i)}$ and $\eta_{h(i)}$ are absolute histology random effects (RE) and relative adult vs. pediatric histology random effects, respectively, and are assumed to be distributed according to:

$$\gamma_h \sim N\left(0, \sigma_\gamma^2\right) \text{ for } h \in \{1, \ldots, H\}$$ (henceforth referred to as absolute cross-histology heterogeneity)

$$\eta_h \sim N\left(0, \sigma_\eta^2\right) \text{ for } h \in \{1, \ldots, H\}$$ (henceforth referred to as relative cross-histology heterogeneity)

- and δ is an optional fixed effect correction term to allow for adult response rates to be mean-shifted on the logit scale (henceforth referred to as delta shift)

- We use the following diffuse priors:

$$\mu \sim \text{Normal}(\text{logit}(0.3), 10^2)$$

$$\delta \sim \text{Normal}(0, 10^2)$$

$$\sigma_\gamma \sim \text{Half}{-}t_2$$

$$\sigma_\eta \sim \text{Half}{-}t_2$$

Demonstration via Simulation: Description (1 / 3)

- We demonstrate the modelling approach using 100 simulated datasets with $n_P = 60$ pediatric patients and $n_A = 180$ adult patients distributed with equal probability across $H = 8$ distinct histologies

- We consider 3 different parameter scenarios:

1. Negligible absolute and relative cross-histology heterogeneity and lower response rates among adults:
   - $\mu = 0, \delta = -1, \sigma_\gamma = 0.01, \sigma_\eta = 0.01$

2. Absolute cross-histology heterogeneity only, and lower response rates among adults:
   - $\mu = 0, \delta = -1, \sigma_\gamma = 0.8, \sigma_\eta = 0.01$

3. Both absolute and relative cross-histology heterogeneity, and lower response rates among adults:
   - $\mu = 0, \delta = -1, \sigma_\gamma = 0.8, \sigma_\eta = 0.8$

- Histologies 1 through 8 are ordered by increasing prospect of response

Demonstration via Simulation: Description (2 / 3)

- We compute and compare 95% credible intervals (CrI) for pediatric response rates under 3 models:

1. Ignore adult data and fit a model with only absolute cross-histology heterogeneity random effect ("1-RE no pooling")
   - Assumes that histology is prognostic for response but allows for partial pooling of information across histologies based on the amount of observed heterogeneity in outcomes between histologies

2. Absolute cross-histology random effect with delta shift ("1-RE with delta")
   - As in model (1), assumes that histology is prognostic for the outcome and allows for partial pooling across histologies
   - Assumes that histology-specific prognosis is similar for adults and pediatric patients after accounting for a higher/lower response in adult patients relative to children that is common across histologies (i.e. age is similarly prognostic for response regardless of histology)

3. Both absolute and relative cross-histology random effects with delta shift ("2-RE with delta")
   - As in model (1) and (2), assumes that histology is prognostic for the outcome and allows for partial pooling across histologies
   - As in model (2), allows for adult patients to have higher/lower response than similar pediatric patients
   - Relaxes model (2) assumption that histology-specific responses are similar between adult and pediatric patients; will attenuate the amount of borrowing from adult patients if within-histology response tends to differ between adult and pediatric patients apart from delta shift

Demonstration via Simulation: Description (3 / 3)

- We would expect:
  - The “1-RE with delta” model to yield more precise estimates compared to the “1-RE no pooling” model due to the addition of information borrowing from adult patients
  - The “2-RE with delta” model to yield less precise estimates than the “1-RE with delta model” but reduce the risk of false positives when borrowing from adult populations is unwarranted (i.e. when histology specific responses differ notably between adult and pediatric patients apart from the delta shift)

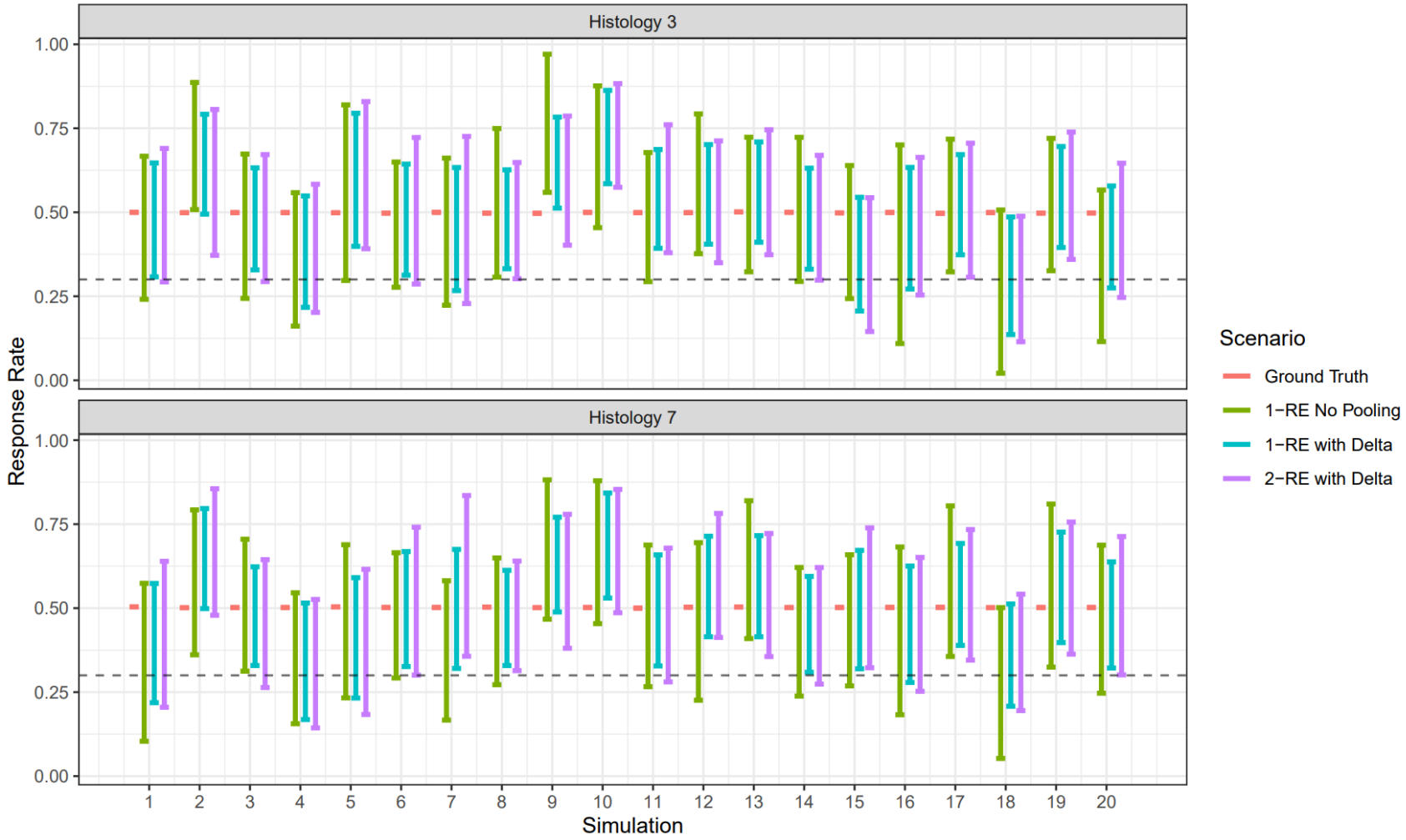
- For each model, we consider how often we conclude efficacy under a decision rule that the lower bound of the 95% CrI must exceed a response rate of 30%. We also plot the 95% CrIs vs. the true response rate (“ground truth”) for the first 20 simulated datasets.

- 95% CrIs were computed via Markov chain Monte Carlo (MCMC) implemented using Stan. MCMC convergence was assessed via $\hat{R}$ statistics [7]

- Note: for this presentation we focus on type-II error implications rather than type-I

Scenario 1: Negligible Absolute and Relative Cross-histology Heterogeneity and Lower Response Rates among Adults

- Figure plots 95% CrIs for the pediatric response rate under different modelling approaches vs. the true response rate (ground truth) for the first 20 simulated datasets
- "1-RE no pooling" estimates are imprecise
- "1-RE with delta" model tends to have narrower CrIs than "1-RE no pooling" approach and good capture of true response. "2-RE with delta" model has slightly wider CrIs than "1-RE with delta" model.



Scenario 1: Negligible Absolute and Relative Cross-histology Heterogeneity and Lower Response Rates among Adults

- For histology 3, conclude efficacy for
  - 41 / 100 true cases (response > 30%) for “1-RE no pooling” approach,
  - 67 / 100 for “1-RE with delta” model, and
  - 46 / 100 for “2-RE with delta” model

- For histology 7, conclude efficacy for
  - 33 / 100 true cases (response > 30%) for “1-RE no pooling” approach,
  - 66 / 100 for “1-RE with delta” model, and
  - 49 / 100 for “2-RE with delta” model

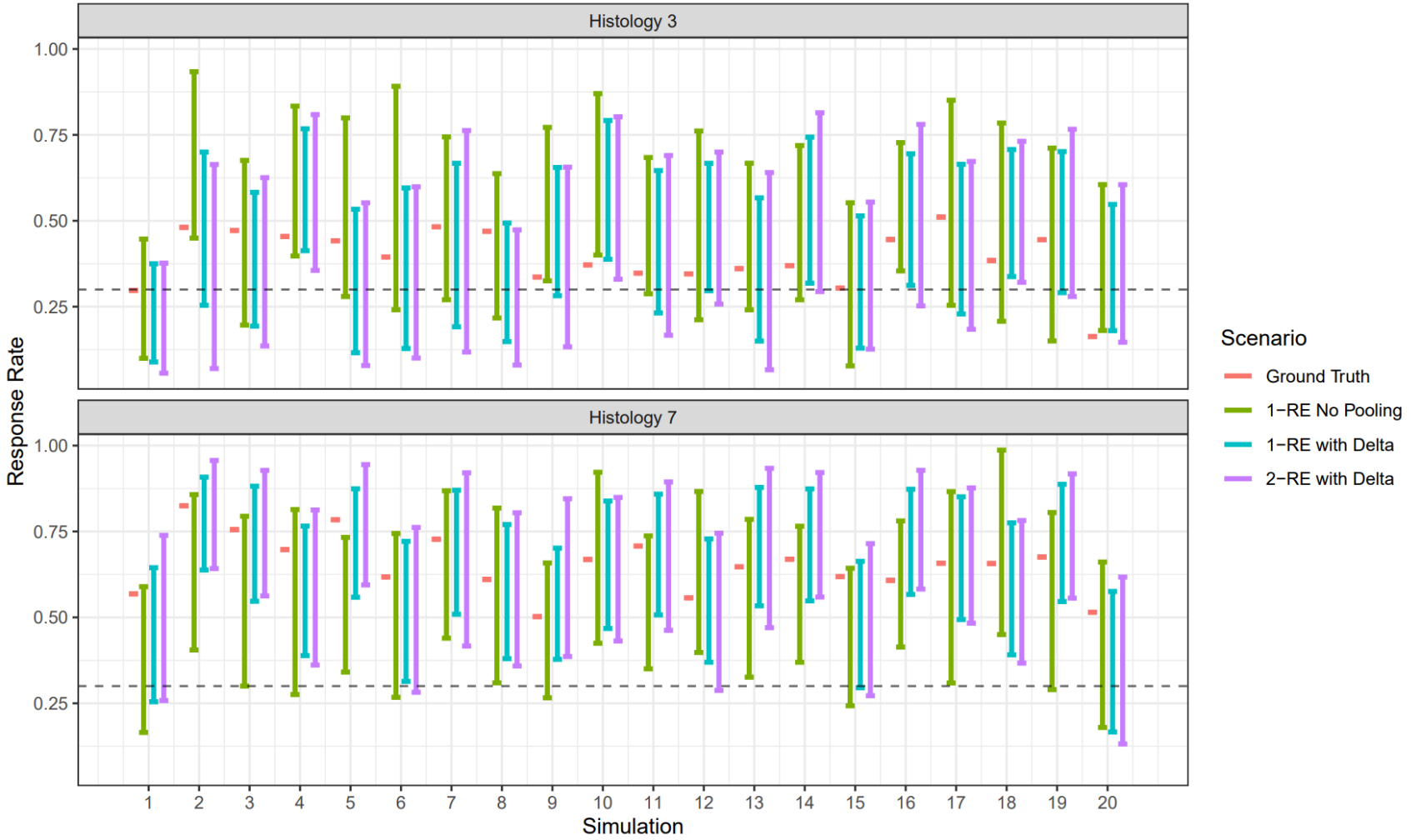
Histology	Model	Meets Efficacy Condition	
		N	%
1	1-RE No Pooling	38 / 100	38
1	1-RE with Delta	65 / 100	65
1	2-RE with Delta	50 / 100	50
2	1-RE No Pooling	40 / 100	40
2	1-RE with Delta	69 / 100	69
2	2-RE with Delta	54 / 100	54
3	1-RE No Pooling	41 / 100	41
3	1-RE with Delta	67 / 100	67
3	2-RE with Delta	46 / 100	46
4	1-RE No Pooling	43 / 100	43
4	1-RE with Delta	66 / 100	66
4	2-RE with Delta	50 / 100	50
5	1-RE No Pooling	42 / 100	42
5	1-RE with Delta	65 / 100	65
5	2-RE with Delta	41 / 100	41
6	1-RE No Pooling	37 / 100	37
6	1-RE with Delta	62 / 100	62
6	2-RE with Delta	43 / 100	43
7	1-RE No Pooling	33 / 100	33
7	1-RE with Delta	66 / 100	66
7	2-RE with Delta	49 / 100	49
8	1-RE No Pooling	47 / 100	47
8	1-RE with Delta	72 / 100	72
8	2-RE with Delta	49 / 100	49

Scenario 2: Absolute Cross-histology Heterogeneity Only and Lower Response Rates among Adults

- Figure plots 95% CrIs for the pediatric response rate under different modelling approaches vs. the true response rate (ground truth) for the first 20 simulated datasets

- "1-RE no pooling" estimates are imprecise—unsurprising as higher absolute cross-histology heterogeneity means less partial pooling from other histologies

- "1-RE with delta" model tends to have much narrower CrIs than "1-RE no pooling" approach and good capture of true response. "2-RE with delta" model has slightly wider CrIs



Scenario 2: Absolute Cross-histology Heterogeneity Only and Lower Response Rates among Adults

- For histology 3, conclude efficacy for
  - 18 / 100 true cases (response > 30%) for “1-RE no pooling” approach,
  - 21 / 100 for “1-RE with delta” model, and
  - 16 / 100 for “2-RE with Delta” model

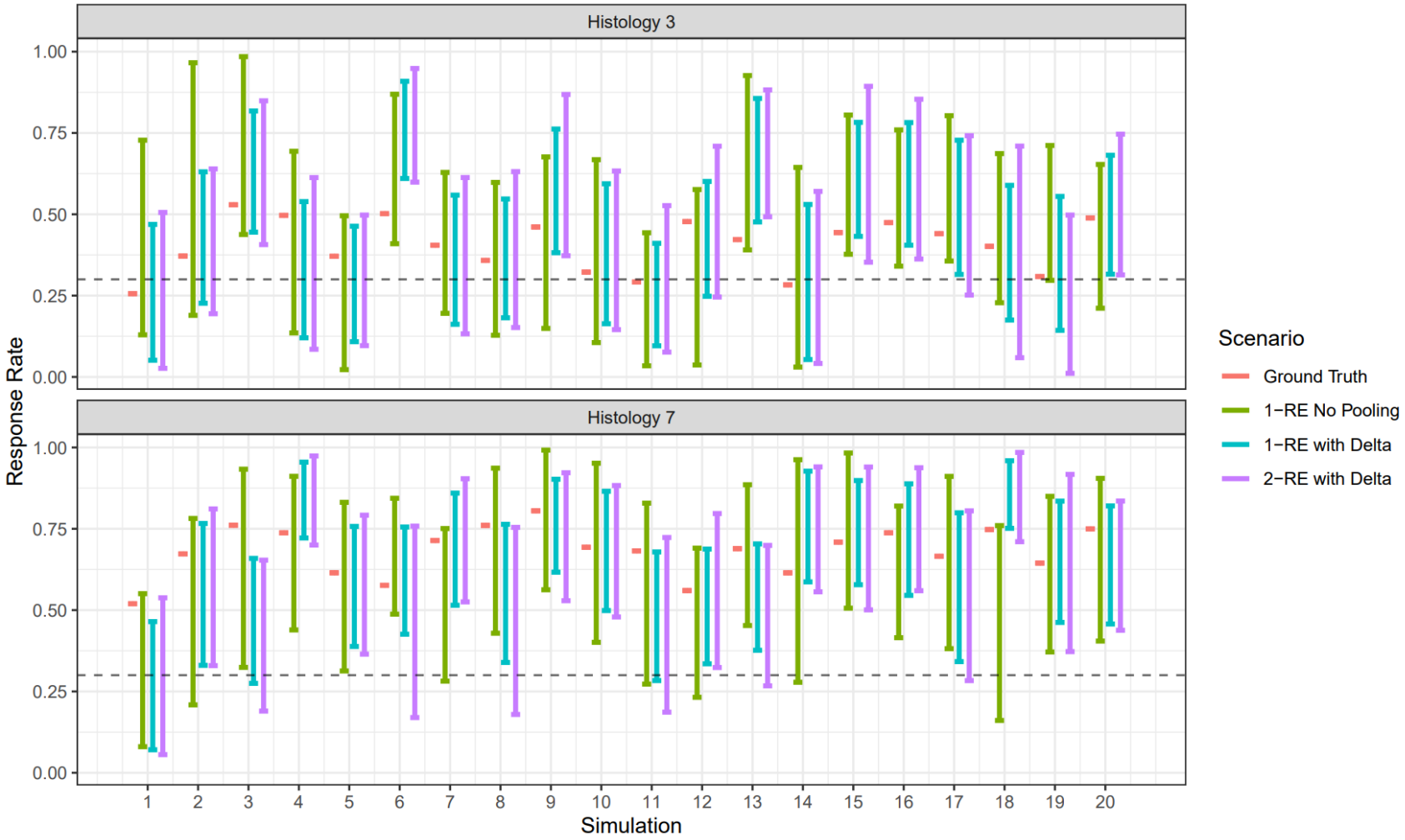
- For histology 7, conclude efficacy for
  - 56 / 100 true cases (response > 30%) for “1-RE no pooling” approach,
  - 78 / 100 for “1-RE with delta” model, and
  - 69 / 100 for “2-RE with delta” model

Histology	Model	Meets Efficacy Condition	
		N	%
1	1-RE No Pooling	0 / 15	0
1	1-RE with Delta	2 / 15	13.3
1	2-RE with Delta	1 / 15	6.7
2	1-RE No Pooling	10 / 62	16.1
2	1-RE with Delta	9 / 62	14.5
2	2-RE with Delta	5 / 62	8.1
3	1-RE No Pooling	18 / 94	19.1
3	1-RE with Delta	21 / 94	22.3
3	2-RE with Delta	16 / 94	17
4	1-RE No Pooling	26 / 100	26
4	1-RE with Delta	35 / 100	35
4	2-RE with Delta	24 / 100	24
5	1-RE No Pooling	35 / 100	35
5	1-RE with Delta	49 / 100	49
5	2-RE with Delta	37 / 100	37
6	1-RE No Pooling	46 / 100	46
6	1-RE with Delta	62 / 100	62
6	2-RE with Delta	53 / 100	53
7	1-RE No Pooling	56 / 100	56
7	1-RE with Delta	78 / 100	78
7	2-RE with Delta	69 / 100	69
8	1-RE No Pooling	72 / 100	72
8	1-RE with Delta	89 / 100	89
8	2-RE with Delta	91 / 100	91

Scenario 3: Both Absolute and Relative Cross-histology Heterogeneity, and Lower Response Rates among Adults

- Figure plots 95% CrIs for the pediatric response rate under different modelling approaches vs. the true response rate (ground truth) for the first 20 simulated datasets

- Differences between the “1-RE with Delta” and “2-RE with Delta” models are more pronounced as the “2-RE with Delta” model is able to account for the addition of relative cross-histology heterogeneity, tempering the amount of borrowing from the adult data



Scenario 3: Both Absolute and Relative Cross-histology Heterogeneity, and Lower Response Rates among Adults

- For histology 3, conclude efficacy for
  - 18 / 100 true cases (response > 30%) for “1-RE no pooling” approach,
  - 30 / 100 for “1-RE with Delta” model, and
  - 24 / 100 for “2-RE with Delta” model

- For histology 7, conclude efficacy for
  - 53 / 100 true cases (response > 30%) for “1-RE no pooling” approach,
  - 74 / 100 for “1-RE with Delta” model, and
  - 61 / 100 for “2-RE with Delta” model

Histology	Model	Meets Efficacy Condition	
		N	%
1	1-RE No Pooling	3 / 22	13.6
1	1-RE with Delta	5 / 22	22.7
1	2-RE with Delta	3 / 22	13.6
2	1-RE No Pooling	6 / 68	8.8
2	1-RE with Delta	16 / 68	23.5
2	2-RE with Delta	12 / 68	17.6
3	1-RE No Pooling	18 / 83	21.7
3	1-RE with Delta	30 / 83	36.1
3	2-RE with Delta	24 / 83	28.9
4	1-RE No Pooling	20 / 99	20.2
4	1-RE with Delta	40 / 99	40.4
4	2-RE with Delta	33 / 99	33.3
5	1-RE No Pooling	38 / 99	38.4
5	1-RE with Delta	48 / 99	48.5
5	2-RE with Delta	40 / 99	40.4
6	1-RE No Pooling	57 / 100	57
6	1-RE with Delta	66 / 100	66
6	2-RE with Delta	52 / 100	52
7	1-RE No Pooling	53 / 100	53
7	1-RE with Delta	74 / 100	74
7	2-RE with Delta	61 / 100	61
8	1-RE No Pooling	74 / 100	74
8	1-RE with Delta	90 / 100	90
8	2-RE with Delta	80 / 100	80

Limitations

- Data limitations present a challenge for reliable model estimation
  - Achieving good MCMC mixing was challenging ($\hat{R}$ >= 1.05 for 0.9% and ESS < 1000 for 4.2% of posterior pediatric response rates across various simulated datasets despite 4 chains of 10,000 iterations with burn-in of 1,000)
  - Potentially very few histologies, making it difficult to estimate σ_γ and σ_η random effect heterogeneity parameters—strong priors may be necessary
  - Potentially very few patients per histology—especially for pediatric patients—can present a further challenge (again, strong priors may be necessary)
  - Complexity of parameterization and choice of priors should take into consideration anticipated data availability

- Likely to be limited ability in practice to control for additional prognostic factors which may be imbalanced without excessive overparameterization issues
  - I.e. borrowing from adult into pediatric population assumes that it is appropriate to borrow from adult populations after accounting for histology random effects and “delta” shift in mean response between pediatric and adult populations—perhaps adult and pediatric patients tend to differ in other prognostic baseline characteristics?

- Various model configurations require potentially strong modelling assumptions to facilitate borrowing
  - Are histologies exchangeable? Is the degree of heterogeneity going to be comparable between pediatric and adult populations? Are there omitted confounders that need to be controlled for?

Potential Extensions

- Can potentially accommodate additional pediatric and/or adult basket trials via an additional random effect (with or without a location shift parameter for pediatric vs. adult trials)
  - May also be an option for incorporation of external information from real-world data (RWD) sources (e.g. RWD for patients receiving a drug with a similar target mutation)

- May be worth down-weighting borrowing from adult trial for “outlier” histologies
  - Allowing for distinct borrowing weights from the adult trial for each histology?
  - Potential to use scale mixtures of normal distributions for the random effects to allow for down-weighting of outliers

- Can we make use of RWD for standard of care treatment to formulate informative priors for some parameters (e.g. random effect variance parameters)?

Conclusions

- We present a flexible modelling framework that can allow for partial pooling of information from pediatric and adult basket trials to improve precision of pediatric response rate estimates
  - Modelling approach can account for a variety of structural assumptions to facilitate partial borrowing

- Care should be taken in deciding on the exact model specification due to anticipated data limitations and risk of overparameterization
  - Model reliability a major concern when data has few histologies with which to estimate heterogeneity parameters and/or few patients per histology
    - Strong priors may be necessary for some parameters!
  - Structural modelling assumptions should be made in consultation with clinical domain experts and probabilistic sensitivity analyses (including prior sensitivity analyses) should be considered where appropriate
  - Practitioners should consider running simulations to see how precision and bias could be impacted under various borrowing strategies, sample sizes, number of histologies, etc. based on their application at hand

References

[1] Park JJ, Siden E, Zoratti MJ, Dron L, Harari O, Singer J, Lester RT, Thorlund K, Mills EJ. Systematic review of basket trials, umbrella trials, and platform trials: a landscape analysis of master protocols. Trials. 2019 Dec;20(1):1-0.

[2] Thall PF, Wathen JK, Bekele BN, Champlin RE, Baker LH, Benjamin RS. Hierarchical Bayesian approaches to phase II trials in diseases with multiple subtypes. Statistics in medicine. 2003 Mar 15;22(5):763-80.

[3] 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 Feb;41(2):165-78.

[4] Food and Drug Administration. BLA 125370/s-064 and BLA 761043/s-007 multi-disciplinary review and evaluation benlysta®(belimumab) for intravenous infusion in children 5 to 17 years of age with SLE. 2021. https://www.fda.gov/media/127912/download.

[5] Viele K, Berry S, Neuenschwander B, Amzal B, Chen F, Enas N, Hobbs B, Ibrahim JG, Kinnersley N, Lindborg S, Micallef S. Use of historical control data for assessing treatment effects in clinical trials. Pharmaceutical statistics. 2014 Jan;13(1):41-54.

[6] Best N, Price RG, Pouliquen IJ, Keene ON. Assessing efficacy in important subgroups in confirmatory trials: An example using Bayesian dynamic borrowing. Pharmaceutical statistics. 2021 May;20(3):551-62.

[7] Vehtari A, Gelman A, Simpson D, Carpenter B, Bürkner PC. Rank-normalization, folding, and localization: An improved $\hat{R}$ for assessing convergence of MCMC (with discussion). Bayesian analysis. 2021.

Thank you!

Contact: Emma Mackay,
emma.mackay@cytel.com

Cytel