Bayesian Hierarchical Models & Information Borrowing

Kyle Mann

Nov 17, 2023

Non-Hierarchical Models

- Beta-Binomial Model

$$y_i \sim Bern(p)$$
$$p \sim Beta(\alpha, \beta)$$

α and β are constants

Exchangeability

- Validity of beta-binomial model relies on exchangeability of the data points.
- Exchangeability means the joint distribution of $\boldsymbol{y}$ is invariant to permutation of the indices. Does not require independence! A multivariate normal distribution is exchangeable if it has independent or compound symmetric covariance, but not if it has autoregressive or block diagonal covariance.
- Partial/conditional exchangeability:
  - Partial: There is a grouping subdividing $\boldsymbol{y}$
  - Conditional: y_i not exchangeable but (y_i, x_i) is exchangeable
- Beta-binomial model not valid if we have partial or conditional exchangeability! What to do?

Partial or Conditional Exchangeability

- Approach 1: Adding covariates. Let X be a categorical covariate.

Bayesian Logistic regression:

$$y_i \sim Bern(p_i)$$

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_i$$

$$\beta_0 \sim N(0, 10000)$$

$$\beta_1 \sim N(0, 10000)$$

Rat Tumor Data

y	N
0	20
0	20
0	20
0	20
0	20
0	20
0	20
0	19
0	19
0	19
0	19
0	18
0	18
0	17
1	20
1	20
1	20
1	20
1	19
1	19
1	18
1	18
3	27
2	25
2	24
2	23
2	20
2	20
2	20
2	20
2	20
2	20
2	20
1	10
5	49
2	19
5	46
2	17
7	49
7	47
3	20
3	20
2	13
9	48
10	50
4	20
4	20
4	20
4	20
4	20
4	20
10	48
4	19
4	19
4	19
5	22
11	46
12	49
5	20
5	20
6	23
5	19
6	22
6	20
6	20
6	20
16	52
15	46
15	47
9	24
4	14

We could:

A) Pool the data and use a simple beta-binomial model.

B) Analyze each experiment separately with a beta-binomial model.

C) Include an indicator variable for each experiment as covariates in a Bayesian logistic regression model. This allows partition of variance but does not borrow information about the mean.

D) Fit a hierarchical model.

Hierarchical Model

We could:

A) Pool the data and use a simple beta-binomial model.

B) Analyze each experiment separately with a beta-binomial model.

C) Include an indicator variable for each experiment as covariates in a Bayesian logistic regression model. This allows partition of variance but does not borrow information about the mean.

D) Fit a hierarchical model.

Partial or Conditional Exchangeability

- Approach 2: Hierarchical model

$$y_{ij} \sim Bern(p_i)$$

$$p_i \sim Beta(\alpha, \beta)$$

$$p(\alpha, \beta) \propto (\alpha + \beta)^{-5/2}$$

$$y_i \sim Binom(n_i, p_i)$$

$$p_i \sim Beta(\alpha, \beta)\beta_0 + \beta_1 x_i$$

$$p(\alpha, \beta) \propto (\alpha + \beta)^{-5/2}$$

Partial or Conditional Exchangeability

- Approach 2: Hierarchical model

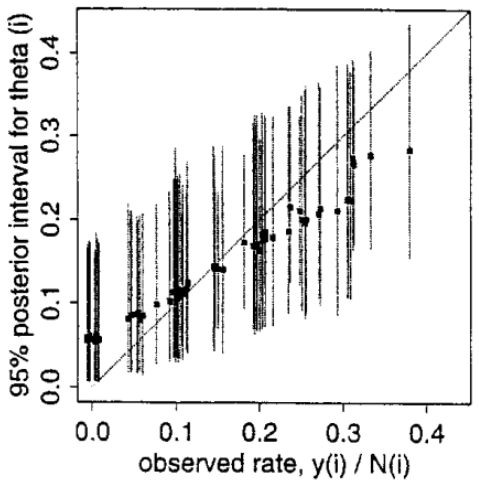
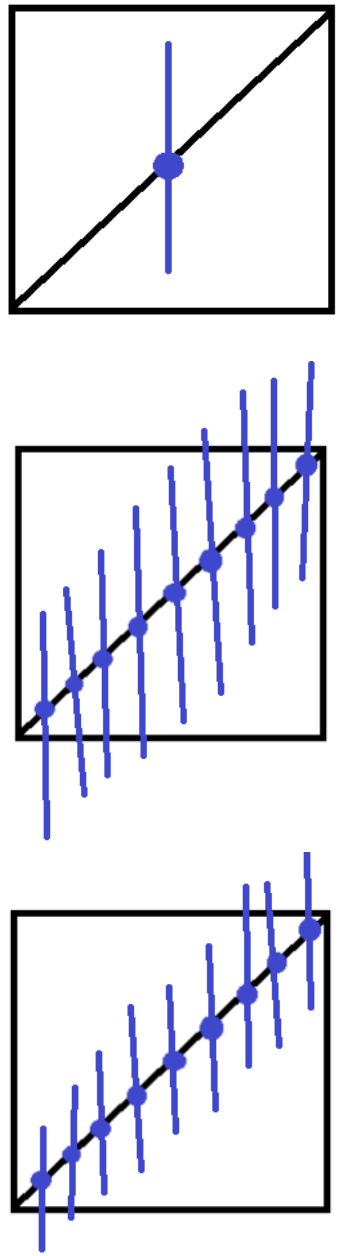


Figure 5.4 *Posterior medians and 95% intervals of rat tumor rates, θ_j (plotted vs. observed tumor rates y_j/n_j), based on simulations from the joint posterior distribution. The 45° line corresponds to the unpooled estimates, $\hat{\theta}_i = y_i/n_i$. The horizontal positions of the line have been jittered to reduce overlap.*



Normal Hierarchical Model

$$y_{ij}|\theta_j \sim N\left(\theta_j, \sigma^2\right)$$

How to set the prior?

Consider the weighted combination: $\hat{\theta}_j = \lambda_j \bar{y}_{*j} + (1 - \lambda_i)\bar{y}_{**}$

- $\hat{\theta}_j = \bar{y}_{*j}$ is the posterior mean if θ_j have independent uniform priors on R

- $\hat{\theta}_j = \bar{y}_{**}$ is the posterior mean if θ_j are restricted to be equal with a uniform prior on R

- $\hat{\theta}_j = \lambda_j \bar{y}_{*j} + (1 - \lambda_i)\bar{y}_{**}$ is the posterior mean if θ_j have iid normal prior densities

BDA uses: $\theta_i \sim N(\mu, \tau^2)$ and $p(\mu, |\tau) \propto 1$

MANDALA Trial and Pediatric Extrapolation

- Airsupra is a combination of salbutamol (a β_2-adrenergic agonist), and budesonide (a corticosteroid), intended to treat asthma.
- Three Phase 3 trials were carried out, MANDALA (n=3132), DENALI (n=1001), and TYREE (n=60). MANDALA focused on asthma attacks and DENALI focused on bronchodilation. MANDALA & DENALI both had positive primary results and led to the approval of Airsupra adult asthma patients.

MANDALA Trial and Pediatric Extrapolation

- In MANDALA, Airsupra had significant effect in adult population (n=2944). There were fewer children (n=83) and adolescents (n=100).
- There were few asthma attacks in the child/adolescent populations, and so the estimates had high uncertainty for these populations.
- FDA recommended Bond Avillion conduct post-hoc Bayesian information-borrowing analysis.
- “Furthermore, recent FDA guidance states that when adult data are available in conditions existing in adults, adolescents, and children, evidence of clinical benefit from the drug in adults can provide support for the prospect of direct benefit in pediatrics. This approach is reasonable for BDA MDI because of similarities in clinical and pharmacologic aspects of asthma across all age groups”

MANDALA Trial and Pediatric Extrapolation

- FDA & Bond Avillion conducted independent analyses:
- FDA used mixture prior
- Bond Avillion used a hierarchical model:
  - Modeled hazard ratios across dose & age subgroups. Age broken into 4-11, 12-17, 18-64, and 65+. There were 2 dose levels. Y_k is log time to first event.

$$Y_k \sim N(\theta_k, \sigma^2) \text{ for k = 1,2,..., 7}$$

$$\theta_k \sim N(\mu, \tau^2)$$

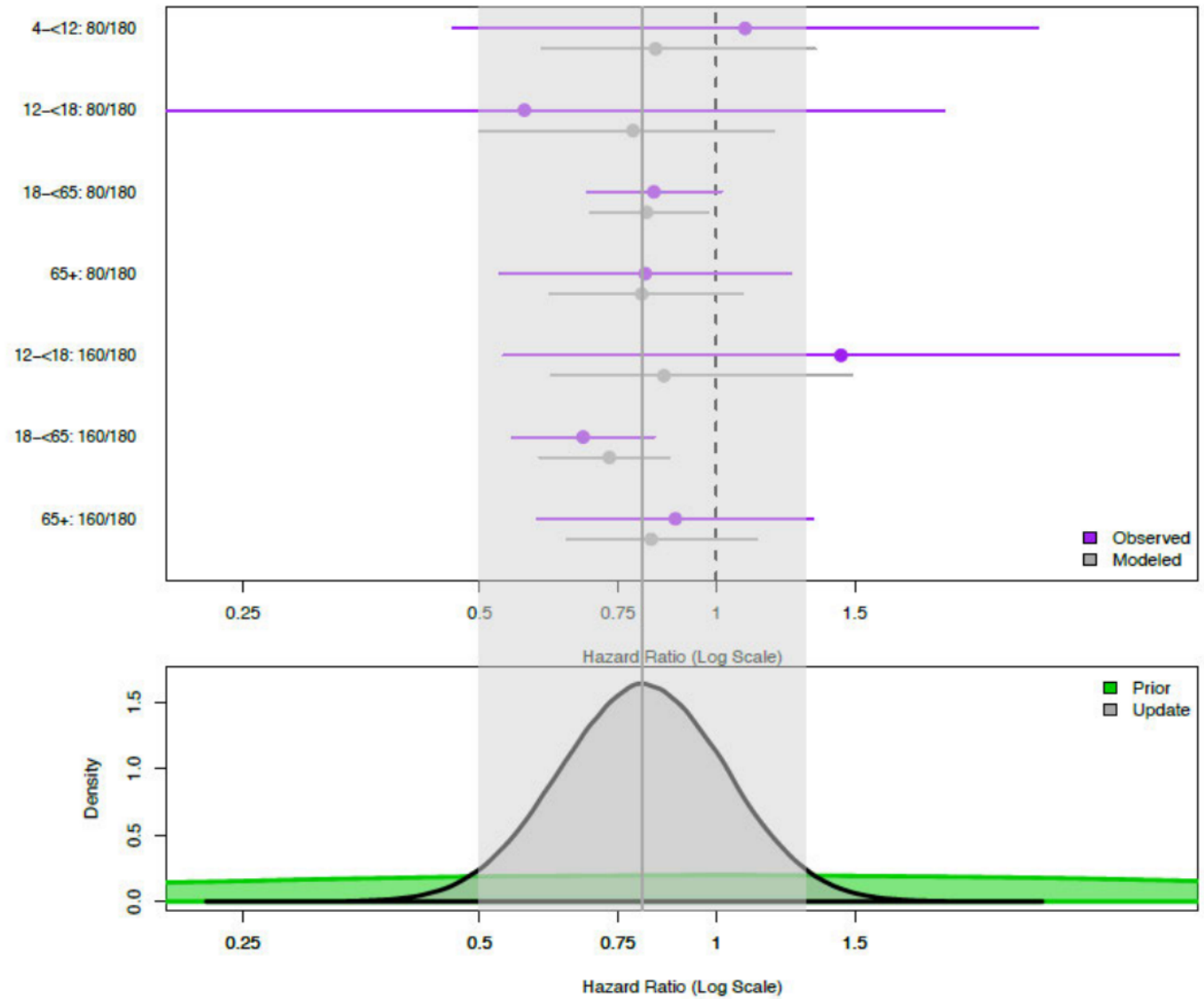
$$\mu \sim N(0, 2^2)$$

$$\tau \sim HalfNormal(2)$$

Sensitivity Analyses

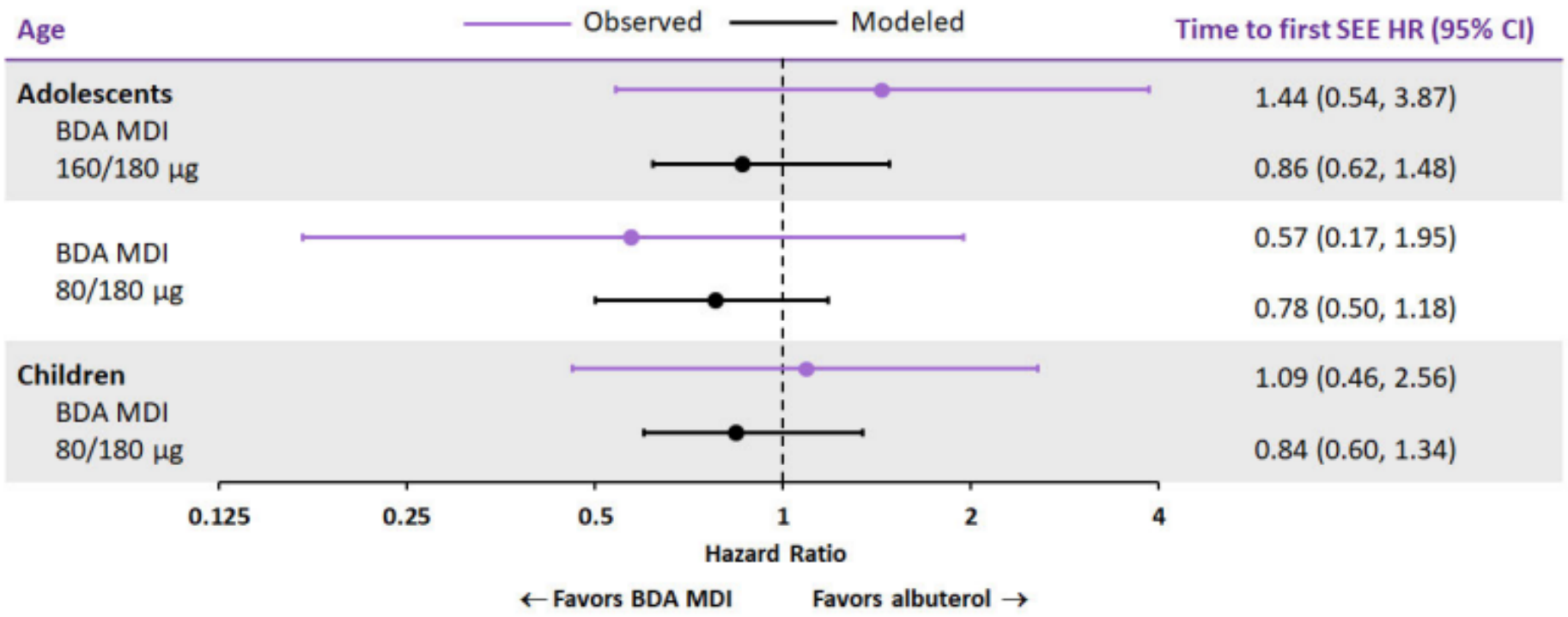
We conducted a sensitivity analysis to understand how estimation for the pediatric population would vary with different assumptions on the prior distributions. This sensitivity analysis focuses on the prior value for τ. This parameter is the explicit measure of the heterogeneity across subgroups and, as demonstrated above, governs the amount of borrowing. If τ is very large, subgroups are allowed to be different and weaker borrowing occurs. If τ is very small, then subgroups are assumed to be very similar and stronger borrowing will occur. While we place a prior on τ, which is then updated by the observed data, the unit of analysis for this update is the number of subgroups, which in our model is 7. Therefore, it is possible that the observed data would not overwhelm the prior and the prior could have influence in the model results.

Figure A1 Bayesian Hierarchical Model Fit Across Age and Dose Subgroups



MANDALA Trial and Pediatric Extrapolation

Figure 21 Bayesian Analysis – MANDALA, Primary Endpoint



MANDALA Trial and Pediatric Extrapolation

Table 24. Borrowing Required to Establish Efficacy of Budesonide High Dose in Adolescents[a]

Bayesian Weight on Adults in Prior	Median HR	95% CrI for HR	Number of Borrowed Adult Events	Percentage of Total Events from Adults[b]
0	1.41	(0.54, 3.68)	0	0.0%
0.25	0.98	(0.58, 3.35)	95	84.8%
0.5	0.78	(0.60, 2.95)	218	92.8%
0.75	0.75	(0.61, 2.36)	334	95.2%
0.9	0.74	(0.61, 1.62)	403	96.0%
0.95	0.74	(0.61, 0.98)	427	96.2%
1	0.73	(0.61, 0.88)	455	96.4%

HR, hazard ratio; CrI, credible interval

[a] From Bayesian robust mixture prior model described in Appendix Section 14.4

[b] Calculated as borrowed adult events ÷ (borrowed adult events + events among children + 1)

Table 25. Borrowing Required to Establish Efficacy of Budesonide Low Dose in Children[a]

Bayesian Weight on Adults in Prior	Median HR	95% CrI for HR	Number of Borrowed Adult Events	Percentage of Total Events from Adults[b]
0	1.08	(0.47, 2.50)	0	0%
0.25	0.86	(0.55, 2.13)	175	88.8%
0.5	0.84	(0.64, 1.79)	313	93.4%
0.75	0.84	(0.69, 1.34)	409	94.9%
0.9	0.83	(0.70, 1.02)	458	95.4%
0.96	0.83	(0.70, 1.00)	478	95.6%
1	0.83	(0.70, 0.99)	494	95.7%

HR, hazard ratio; CrI, credible interval

[a] From Bayesian robust mixture prior model described in Appendix Section 14.4

[b] Calculated as borrowed adult and adolescent events ÷ (borrowed adult and adolescent events + events among children + 1)

MANDALA Trial and Pediatric Extrapolation

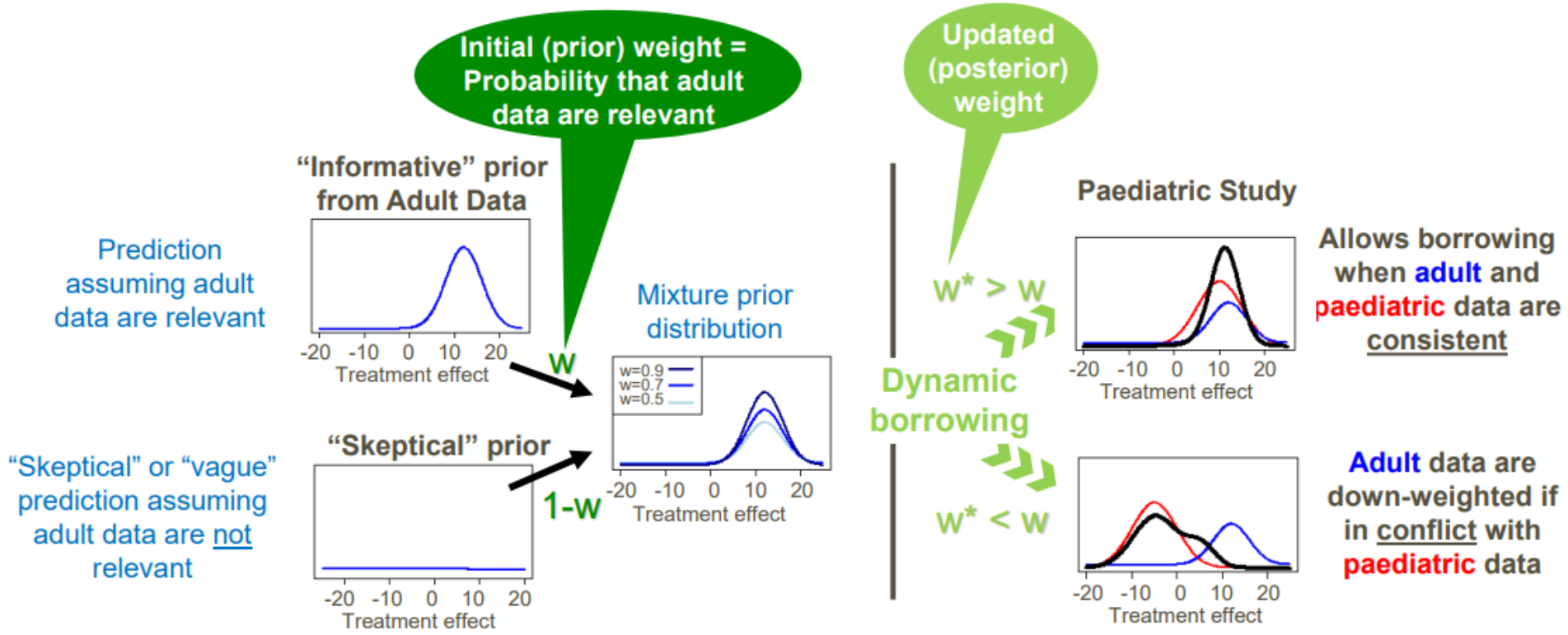
- “In any Bayesian approach based solely on the MANDALA data, to achieve Hazard Ratios with credible intervals <1, which is a high standard of evidence, large amounts of borrowing are needed. This is due to the small sample sizes of both adolescents and children enrolled in the MANDALA study. However, much less borrowing is needed to observe favorable point estimates as demonstrated in this slide. In our Bayesian modeling, borrowing approximately 10% of the available exacerbation events from the overall population was sufficient to observe point estimates of around 15% reduction in risk for adolescents and children, suggesting favorable treatment responses with the proposed BDA MDI doses compared to albuterol.”
- FDA decided against pediatric approval because:
  a) Indication was new, uncertainty about how well the endpoint extrapolates.
  b) Drug was locally acting so extrapolation could not be supported by PK.

Benlysta Approval

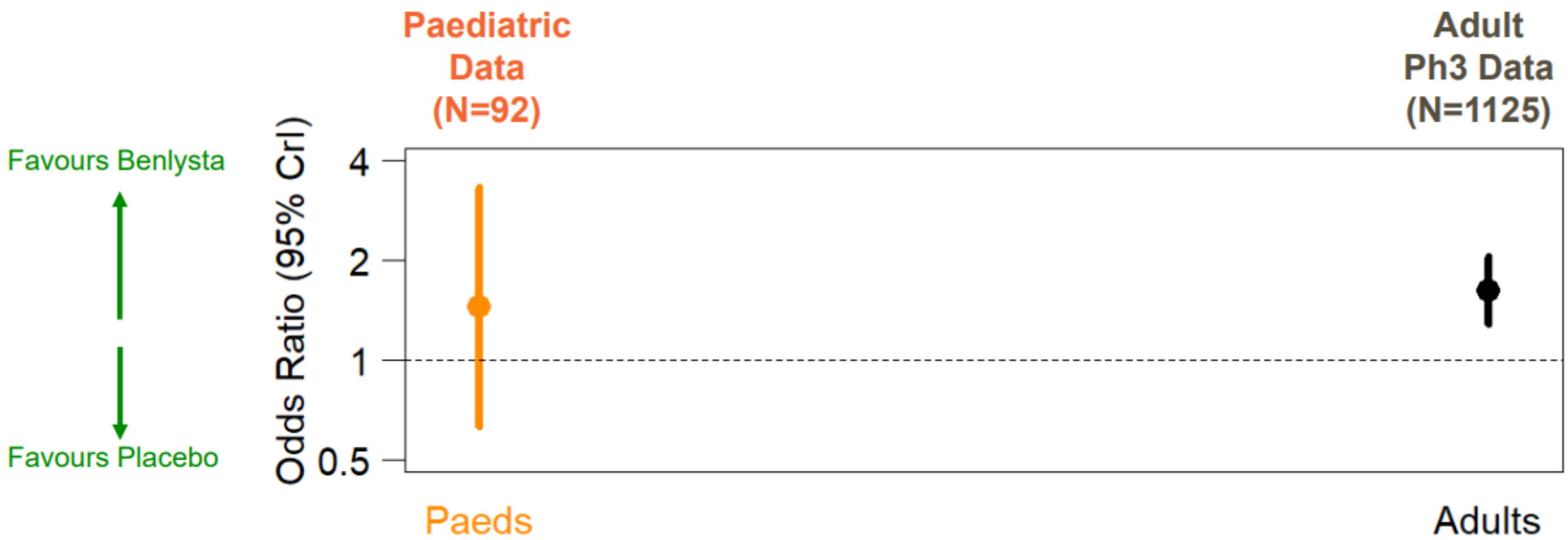
- Benlysta approved for lupus erythematous in 2011. Pediatric study PLUTO started in 2012. Due to rarity of disease, only 93 patients could be recruited. Randomization was 5:1 treatment vs placebo. Planned to “descriptively analyze efficacy and safety”. Not powered for inference.
- FDA requested a post-hoc Bayesian analysis borrowing information from the previous adult trials. The Bayesian analysis served as supporting evidence for an expansion of the indication to include pediatric patients.
- Bayesian dynamic borrowing (Bayesian mixture prior)
- “Tipping point” analysis completed to quantify how much prior belief (range of prior weights) in the applicability of the Adult result it would take in order for the Pediatric study data to look convincing

Benlysta Approval

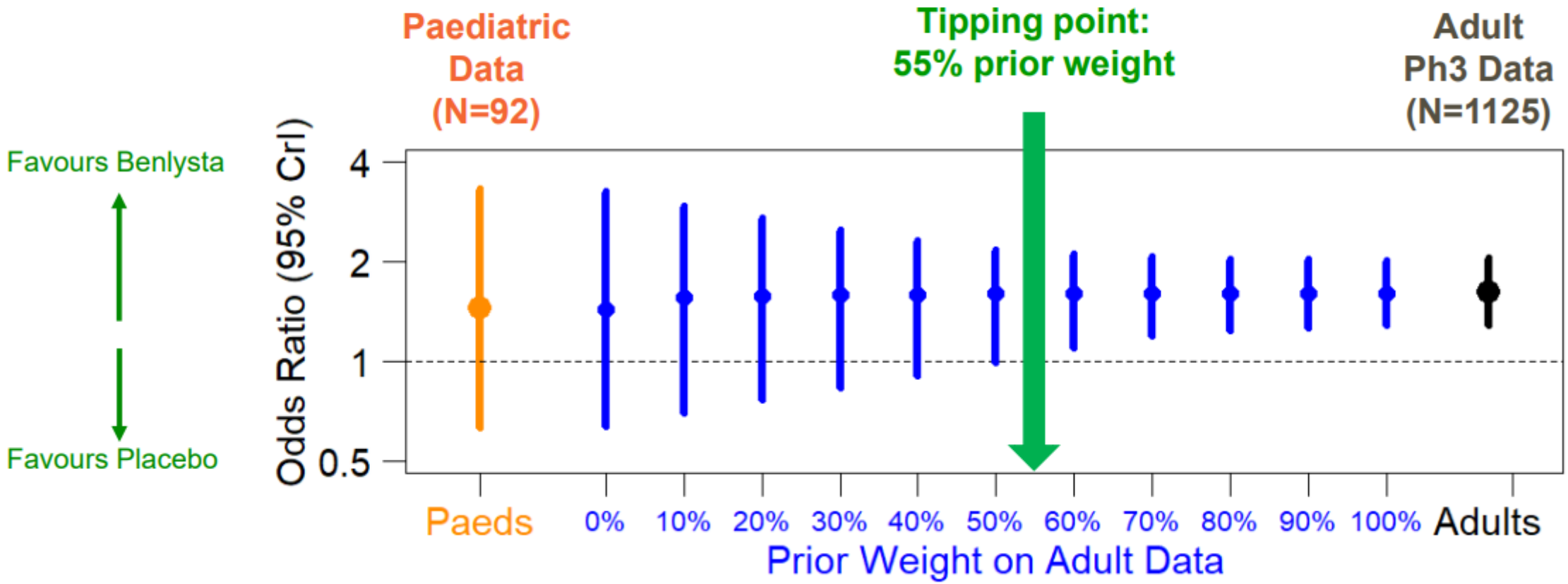
Bayesian Dynamic Mixture priors



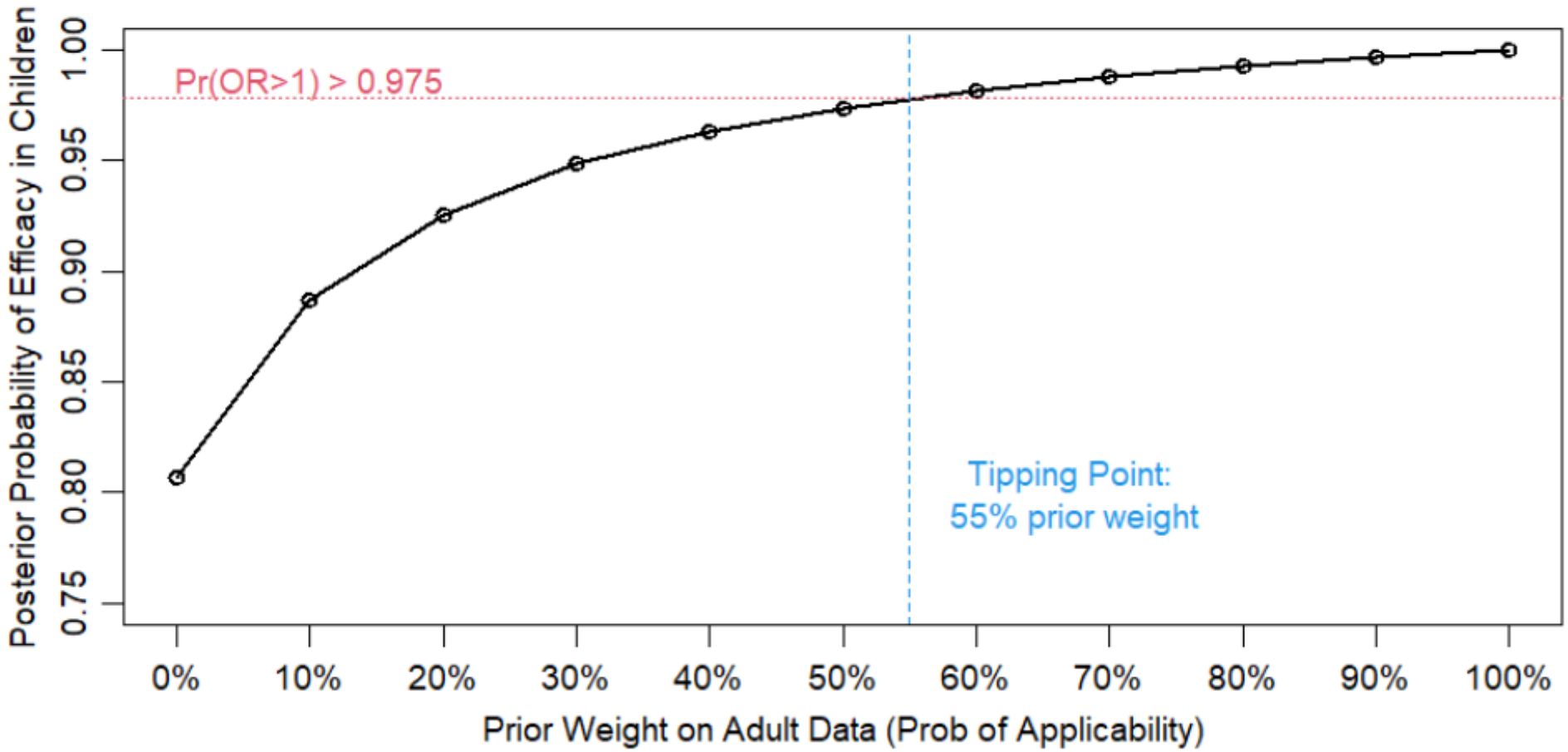
Tipping point Bayesian dynamic mixture prior analysis of Benlysta paediatric study: SRI Responder endpoint



Tipping point Bayesian dynamic mixture prior analysis of Benlysta paediatric study: SRI Responder endpoint



Tipping point Bayesian dynamic mixture prior analysis of Benlysta paediatric study



Celacade – The Importance of Exchangability

- Vasogen developed a medical device Celacade. This device induce apoptosis in a drawn blood sample. The blood was reinjected into the patient and had an anti-inflammatory effect. Intended to treat:
  a) Chronic heart failure
  b) Peripheral artery disease
- In Aug 2005 a Phase 3 peripheral artery disease trial failed.
- In June 2006, ACCLAIM, a Phase 3 chronic heart failure failed its primary endpoint, but reduction of hospitalization in a subpopulation.
- Celacade approved by EMA based on ACCLAIM, but FDA required a further confirmatory trial.

Celacade – ACCLAIM II Trial

- FDA suggested Bayesian design borrowing information from ACCLAIM.
- ACCLAIM II had smaller sample size (n=300-600) than ACCLAIM (n=2400) based on this information-borrowing. Designed by Berry Consultants, a leader in Bayesian clinical trials. Plans announced in Sep 2007.
- In Feb 2008, FDA re-examines, is doubtful that ACCLAIM and ACCLAIM II can recruit comparable populations, and now opposes Bayesian information-borrowing design.
- Vasogen cannot afford larger frequentist design. Lays off 85% of staff and halts ACCLAIM II and Euopean commercialization of Celacade. Stock price tanks.