On the Use of Co-data in Clinical Trials

Satrajit Roychoudhury
Pfizer Inc.
3rd Annual Boston Pharmaceutical Symposium,
Cambridge, MA
May 3, 2019

GLOBAL PRODUCT DEVELOPMENT

Acknowledgement

- Joint work with:
  - Beat Neuenschwander and Heinz Schmidli Novartis Pharma AG

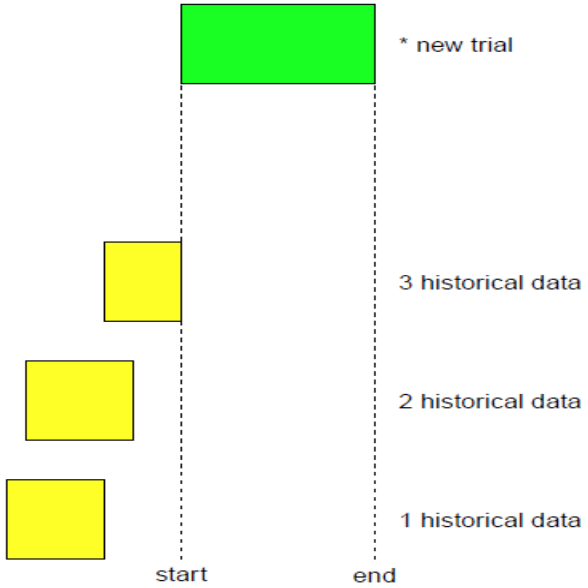
Co-data Approaches in Clinical Trial

- Two recent developments
  - look at the data frequently → adaptive trials
  - look at more data → trials with historical data
- In this talk:
  - we follow the maxim: *more data lead to better decisions*
  - we extend the historical data framework to

    co-data: any relevant complementary/contextual data
  - *co-data can* be *historical* or *concurrent*
- Prospective planning and proper statistical methodology is the key

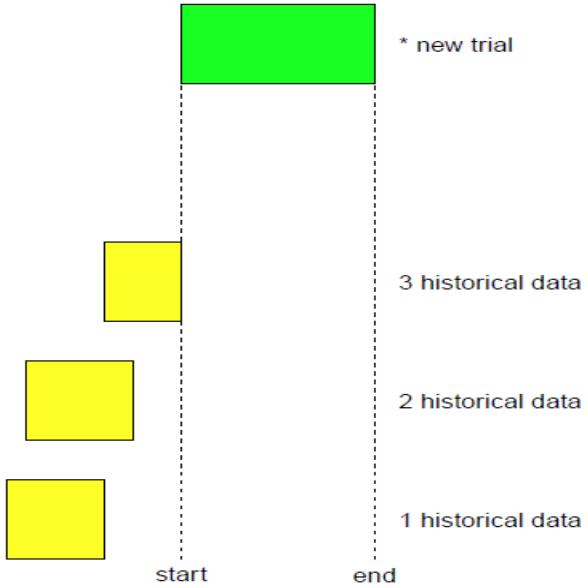
Historical *Co-Data*

- **your (the actual) trial**
- + 3 trials with historical *co-data*
- trial 3 is ongoing...

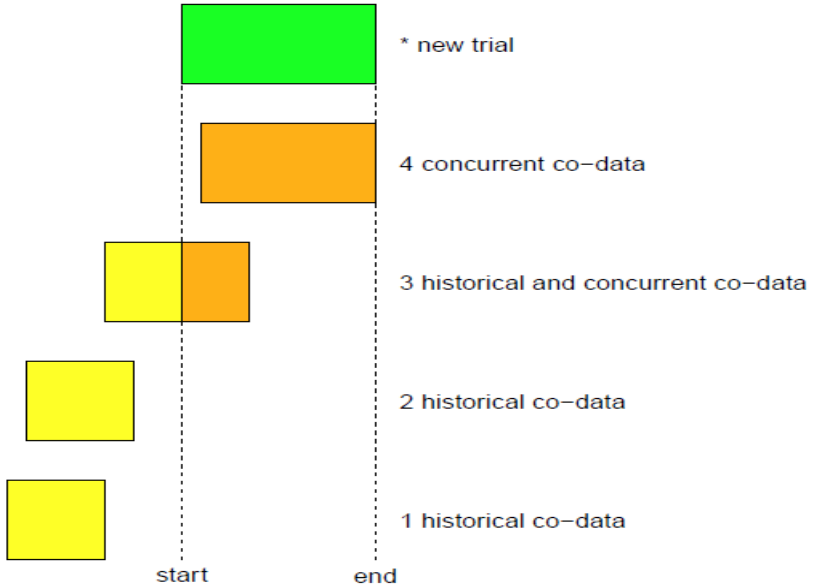


Historical and Concurrent *Co-Data*

- your (the actual trial
- + 3 trials with historical *co-data*...
- trial 3 is ongoing...



- ...trial 3 is ongoing...
- and trial 4 hasn't started yet
- → concurrent *co-data*



Statistical Methodology: Hierarchical model

- Most clinical literature uses two extreme models
  - **No pooling:** Separate inference for each tumor type (stratified analysis) - *"Low power for small sample size situations"*
  - **Complete pooling:** grouping in the data is irrelevant, i.e. imposing restriction that all tumor type effects are same – "optimistic borrowing"
- Bayesian hierarchical modeling is a specific methodology may be used to combine information of different strata.
- **Exchangeable/Hierarchical model** lies between these two extreme cases
  - "Shrinkage": the estimates are pulled towards a common mean

Notable Work

- Full exchangeability of strata parameters is the key assumption for Bayesian hierarchical model discussed in literature:
  - Thall et al. 2003, Chugh et al. 2009, Berry et. al 2013
- General class of nonparametric priors (random partitioning, Polya tree priors etc.) were discussed by
  - Leon-Novelo 2013, Mueller and Mitra 2013
- We propose Tailored exchangeability model based on *meta analytic* approaches
  - borrowing information across similar strata, while avoiding too optimistic borrowing for extreme strata
  - Neuenschwander, Roychoudhury and Schmidli 2016

Meta-Analytic (MA) Approaches

- Two MA approaches
  - ***Meta-Analytic-Predictive* (MAP)** is **prospective**

- At design stage of current trial, perform meta-analysis of *co-data* and obtain distribution of θ_*

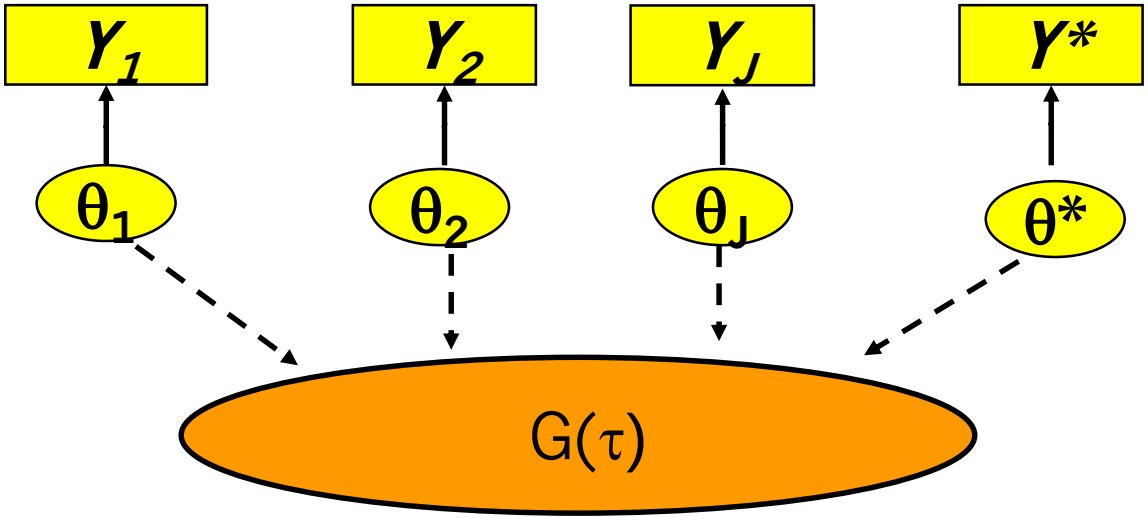
$$\textbf{MAP Prior: } \theta_* \mid Y_1,\ldots Y_C$$

- Combine MAP prior with current trial data Y_* (Bayesian analysis)

- ***Meta-Analytic-Combined (MAC)*** is **retrospective**
  - Perform a meta-analysis of <u>all *co-data* and *current* trial data</u>
  - Parameter of interest: the parameter in the *actual* trial

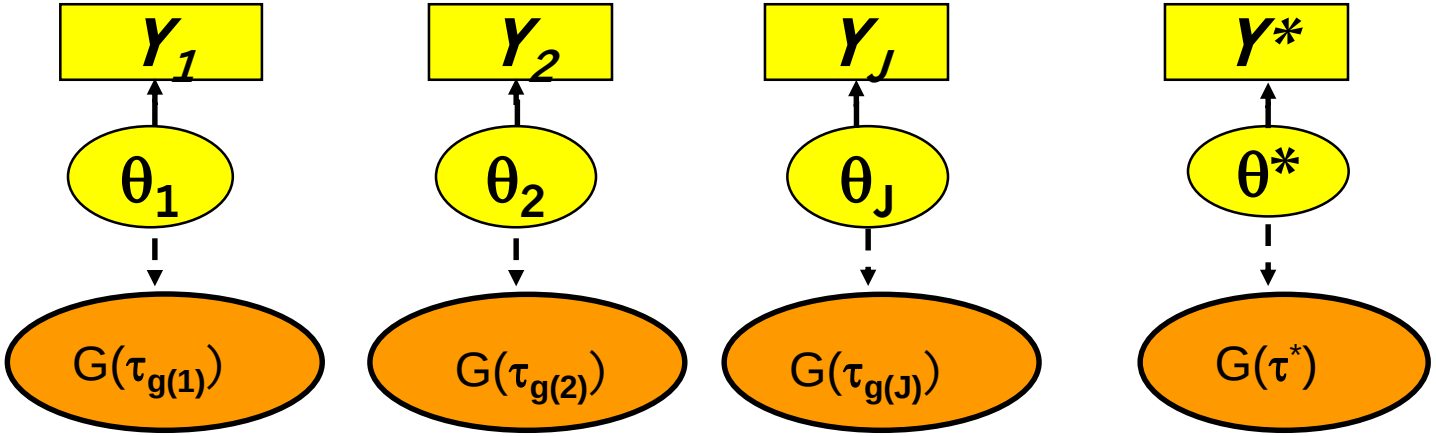
$$\theta_* \mid Y_1,\ldots Y_C, Y_*$$

A Meta-Analytic Framework for Co-Data



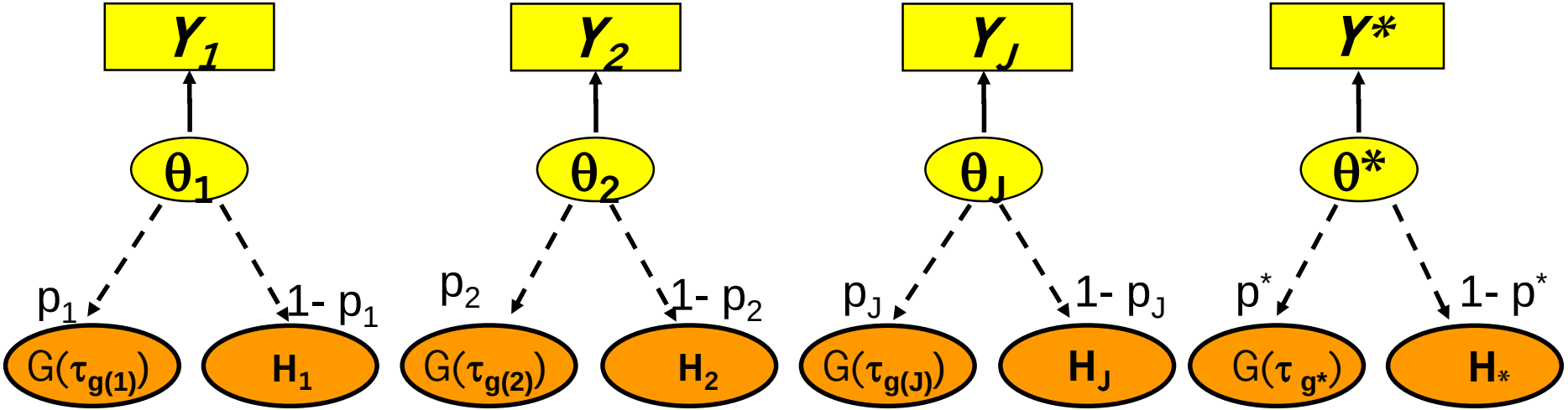
- data (sampling) model $Y_j \mid \theta_j \sim F(\theta_j)$
- parameter model $\theta_1, \ldots, \theta_J, \theta_* \mid \tau \sim G(\tau)$:Exchangeability
- too restrictive if relevance of *co-data* differs
- Possible Extension: adjustment with covariates → Partial exchangeabilty

Flexible Meta-analytic Approach for Co-data



- Extension: $\theta_j \sim G(\tau_{g(j)})$ $g(j) \in \{1,...,G\}$, $j = 1,...J,_*$ Differential discounting
- For example:
  - G=2 for observational and randomized *co-data*
  - note: the larger G, the less information for between-trial sd $\tau_1,..,\tau_G$

A Robust Meta-analytic Approach for Co-data



- **Robustification**: $\theta_j \sim p_j\, G(\tau_{g(j)}) + (1- p_j)\, H_j : g(j) \in \{1,...,G\}, j = 1,...J,_*$
- Allows for nonexchangeable parameters to add robustness

Prior distributions for τ

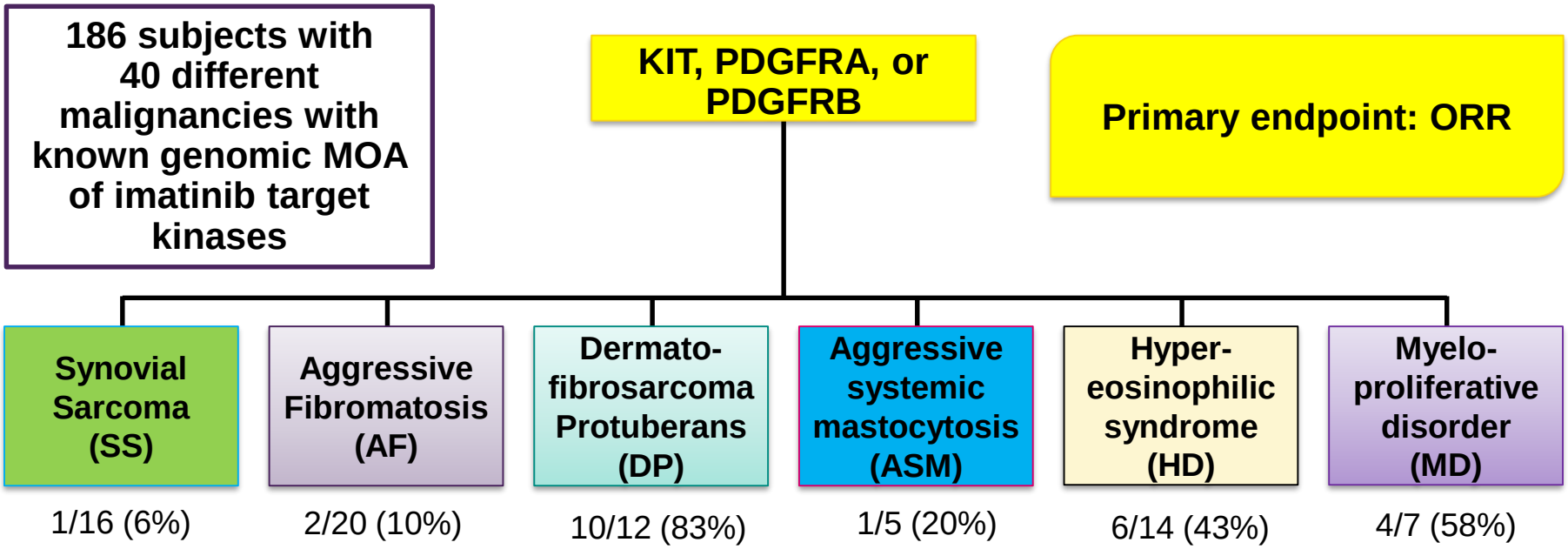
- Since the number of trials (J) is usually small, priors matter
- Recommendations (Spiegelhalter 2004, Gelman 2006)
  - use priors that put most of their probability mass on plausible values
  - example: log-odds scale, $\sigma \approx 2$, half-normal priors with scale 1 and 0.5
    - $\tau \sim$ Half-Normal(1): $(0.03,2.24)_{95\%}$ very small to very large heterogeneity
    - $\tau \sim$ Half-Normal(0.5): $(0.01,1.12)_{95\%}$ very small to large heterogeneity

Weight of Co-Data: Effective Sample Sizes

- Various variance-based approximations to express amount of information of the prior or posterior distribution as an equivalent effective sample size *(ESS)*
  - Malec (2001), Pennello (2008), Morita (2008), N et al (2010)
  - Two-variances approach
    - analysis of interest for θ_*: variance $\boldsymbol{var_*}$ and unknown $\boldsymbol{ESS_*}$
    - simpler analysis: variance $\boldsymbol{var_0}$ and known (!) $\boldsymbol{ESS_0}$; e.g. complete pooling
    - assumption: sample sizes approximately proportional to precisions

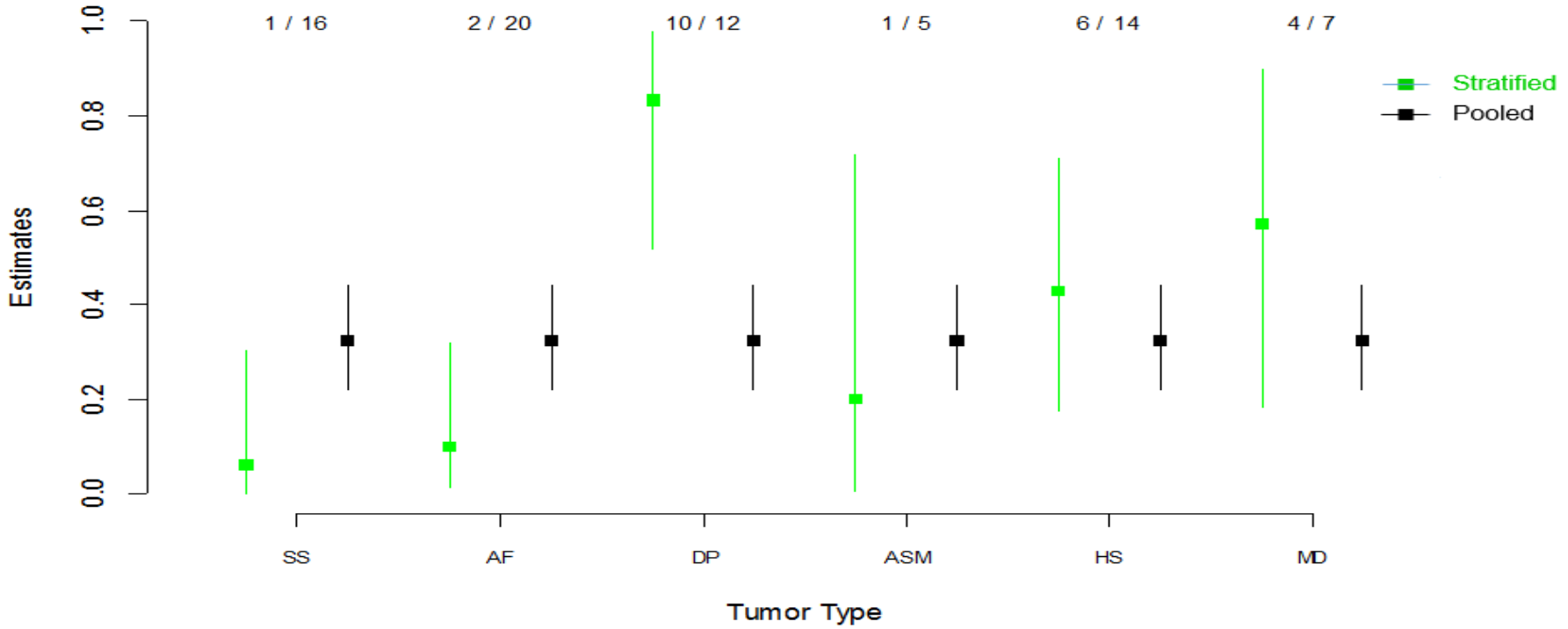
$$ESS_* = ESS_0 \times (var_0 / var_*)$$

Example: Basket Trial of Imatinib



Blumenthal. Innovative trial designs to accelerate the availability of highly effective anti-cancer therapies: an FDA perspective, AACR 2014

Stratified and Pooled Analysis



MAC and Robust MAC Model

- **Data:** n_j = Number of patients and r_j = Number of responder for strata j

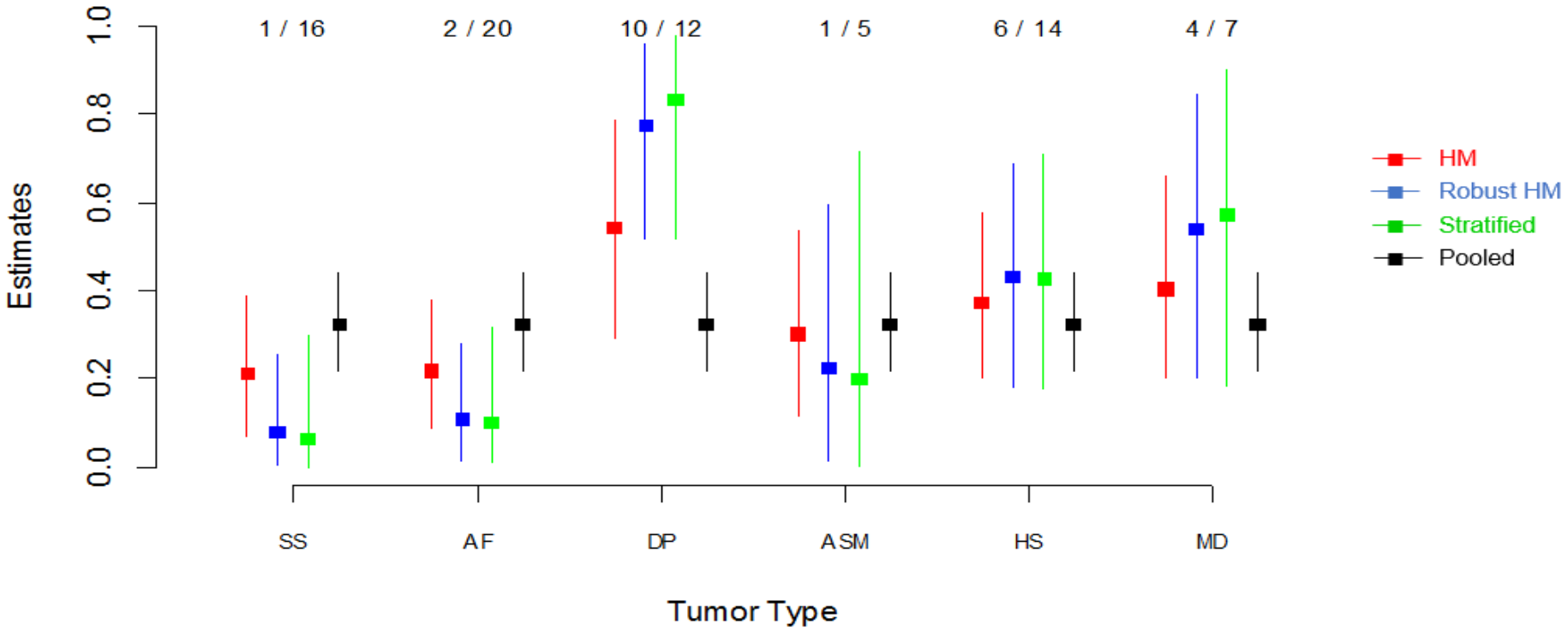
- **Likelihood/sampling model:** $r_j \sim Bin(n_j, \pi_j)$

- **Model:** $\theta_j = \log(\pi_j / 1- \pi_j)$

For each stratum j two possibilities are considered:

  - With probability p_j : $\theta_j \sim N(\mu,\tau^2)$
  - With probability 1- p_j : $\theta_j \sim N(\mathbf{m_J}, \mathbf{v_J})$
  - $p_j = 0$ => MAC or HM
  - For this example we assume $p_j = 0.5$

MAC and Robust MAC Analysis



Example 1: Phase III Interim Analyses

- Two phase III trials *A* and *B* running in parallel
  - endpoint: survival
  - 379 events (n): α=2.5%, 90% power for log-hazard ratio $\theta_A = \log(0.75)$
  - interim analysis when at least 150 deaths occurred in both trials
- Two historical trials
  - a small proof-of-concept trial, and a randomized phase II trial
- Interim decisions
  - based on probability of success *(PoS)*: stop phase III trial if *PoS* < 10% (e.g.)
- *Co-data* analysis with the standard NNHM

$$Y_j \mid \theta_j \sim N(\theta_j, 4/n_j),\ \theta_1,\ldots,\theta_J,\theta_* \mid \mu,\tau \sim N(\mu, \tau^2),\ \mu \sim N(0,4),\ \tau \sim HN(0.5)$$

Stratified Analyses: Estimates and Probability of Success (PoS)

Study	deaths	HR (95%-int)	log(HR) (sd)	pr(HR<1)	PoS
		stratified analyses			
1. Proof-of-concept	8	0.70 (0.18,2.80)	-0.36 (0.71)	0.69	
2. Phase II	85	0.75 (0.49,1.15)	-0.29 (0.22)	0.91	
3. Phase III study A	162	0.83 (0.61,1.13)	-0.19 (0.16)	0.88	0.45
4. Phase III study B	150	0.78 (0.57,1.07)	-0.25 (0.16)	0.94	0.64

- *PoS* calculation requires two components
  - parameter uncertainty at interim: posterior of θ_j
  - conditional power, for example for trial 3, n=379, n_I=162, σ=2

$$CP_3(\theta_3) = \Phi[z_\alpha\sqrt{n/(n-n_I)} - y_3 n_I/(\sigma\sqrt{n-n_I}) - \theta_3\sqrt{n-n_I}/\sigma]$$

  - *PoS* is then the expected (over posterior) conditional power

Co-data Analyses: Estimates and Probability of Success (PoS)

Study	deaths	HR (95%-int)	log(HR) (sd)	pr(HR<1)	PoS
stratified analyses					
1. Proof-of-concept	8	0.70 (0.18,2.80)	-0.36 (0.71)	0.69	
2. Phase II	85	0.75 (0.49,1.15)	-0.29 (0.22)	0.91	
3. Phase III study A	162	0.83 (0.61,1.13)	-0.19 (0.16)	0.88	0.45
4. Phase III study B	150	0.78 (0.57,1.07)	-0.25 (0.16)	0.94	0.64
co-data analysis					
3. Phase III study A	162	0.80 (0.63,1.04)	-0.22 (0.13)	0.95	0.51
4. Phase III study B	150	0.79 (0.61,1.01)	-0.24 (0.13)	0.97	0.65

Co-data analysis

- improves precisions for log-hazard ratios
- *PoS* do not change much

Effective Sample Sizes (ESS)

Study	deaths	HR (95%-int)	log(HR) (sd)	pr(HR<1)	PoS	ESS
stratified analyses						
1. Proof-of-concept	8	0.70 (0.18,2.80)	-0.36 (0.71)	0.69		8
2. Phase II	85	0.75 (0.49,1.15)	-0.29 (0.22)	0.91		85
3. Phase III study A	162	0.83 (0.61,1.13)	-0.19 (0.16)	0.88	0.45	162
4. Phase III study B	150	0.78 (0.57,1.07)	-0.25 (0.16)	0.94	0.64	150
co-data analysis						
3. Phase III study A	162	0.80 (0.63,1.04)	-0.22 (0.13)	0.95	0.51	254
4. Phase III study B	150	0.79 (0.61,1.01)	-0.24 (0.13)	0.97	0.65	252

- Co-data analysis:
  - improves precisions for log-hazard ratios
  - ESS is ≈ 60% larger compared to stratified analyses

Probability of Regulatory Success

- Successful regulatory submission requires both Phase III trials to be positive:
- Probability of regulatory success (*PoRS*)

$$PoRS = \int \mathrm{CP}_3(\theta_3)\mathrm{CP}_4(\theta_4)\, p(\theta_3, \theta_4 | \text{interim data})\, d\theta_3\theta_4$$

Analysis	*PoS* of Trial A	PoS of Trial B	*PoRS*
Full exchangeability	0.51	0.65	0.36
Differential heterogeneity	0.49	0.64	0.34
Exchangeability-nonexchangeability mixture (50-50)	0.49	0.65	0.34

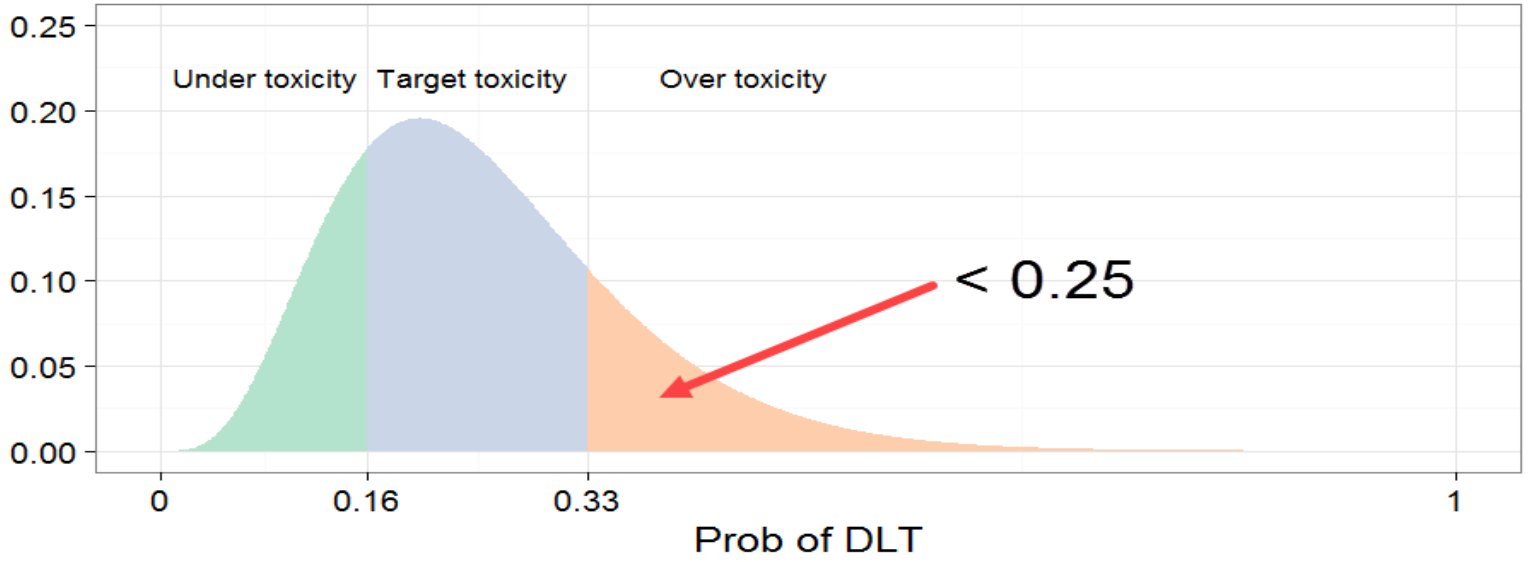
Phase I Combination Trials

- Combination therapies are now popular in Oncology
- Phase I Oncology Trial objectives:
  - Safety and tolerability of patients
  - Find maximum tolerable dose (MTD) or recommended phase II dose
  - data: binary dose-limiting toxicity (DLT) data
- There is no longer one MTD but a many
  - critical to determine the MTD boundary and the set of acceptable doses.
- Overall risk assessment is key
  - Model based approaches: summarize the risk at each dose pair
  - actual decisions use additional information (e.g. efficacy, PK, biomarkers, later cycle AE) to select “best” dose pair(s) for next cohort

Practical Model based Approach for Combination Studies

- Parsimony
  - small number of parameters due to small number of tested dose combinations
- Interpretability
  - easily interpretable parameters for
    - single agent 1 toxicity
    - single agent 2 toxicity
    - interaction
- Continuity
  - if the dose of one compound is 0, the model simplifies to the single-agent model

Escalation with Overdose Criterion (EWOC)



Dose escalation happens when the following condition is satisfied:

- $\Pr(\pi_{ij} > \mathbf{0.33} \mid \text{data}) < 25\%$

Co-data for Phase I Combination Trial (1/3)

- two historical single-agent trials:
  A for agent 1, ongoing
  B for agent 2

1) historical co-data at the start of the AB trial

agent 1 trial A		agent 2 trial B	
3-0	0/3	0-33.3	0/3
4.5-0	0/3	0-50	0/3
6-0	0/6	0-100	0/4
8-0	2/3	0-200	0/9
		0-400	0/15
		0-800	2/20
		0-1120	4/17

Co-data for Phase I Combination Trial (2/3)

- after 3 cohorts of actual trial *AB:* concurrent co-data from trial *A*

1) historical co-data at the start of the AB trial

agent 1 trial A		agent 2 trial B	
3-0	0/3	0-33.3	0/3
4.5-0	0/3	0-50	0/3
6-0	0/6	0-100	0/4
8-0	2/3	0-200	0/9
		0-400	0/15
		0-800	2/20
		0-1120	4/17

2) data after 3 cohorts of AB trial

AB trial		agent 1 trial A	
3-400	0/3	3-0	0/3
3-800	1/3	4.5-0	0/6
6-400	1/3	6-0	0/11
		8-0	2/3

Co-data for Phase I Combination Trial (3/3)

– at end of *AB* trial: co-data from *IIT* combination trial

1) historical co-data at the start of the AB trial

agent 1 trial A		agent 2 trial B	
3-0	0/3	0-33.3	0/3
4.5-0	0/3	0-50	0/3
6-0	0/6	0-100	0/4
8-0	2/3	0-200	0/9
		0-400	0/15
		0-800	2/20
		0-1120	4/17

2) data after 3 cohorts of AB trial

AB trial		agent 1 trial A	
3-400	0/3	3-0	0/3
3-800	1/3	4.5-0	0/6
6-400	1/3	6-0	0/11
		8-0	2/3

3) data at end of AB trial

AB trial		IIT-trial	
3-400	0/3	3-400	0/3
3-800	2/6	3-800	5/7
		4.5-400	0/3
4.5-600	2/10		
6-400	3/10	6-400	0/6
		6-600	2/3

Phase I Trial for Combination Treatment in Cancer

$$\log(\text{odds}(\pi_{1,d_1})) = \log(\alpha_1) + \beta_1 \log(d_1)$$

$$\log(\text{odds}(\pi_{2,d_2})) = \log(\alpha_2) + \beta_2 \log(d_2)$$

$$\pi^0_{12,d_1,d_2} = \pi_{1,d_1} + \pi_{1,d_2} - \pi_{1,d_1}\pi_{2,d_2}$$

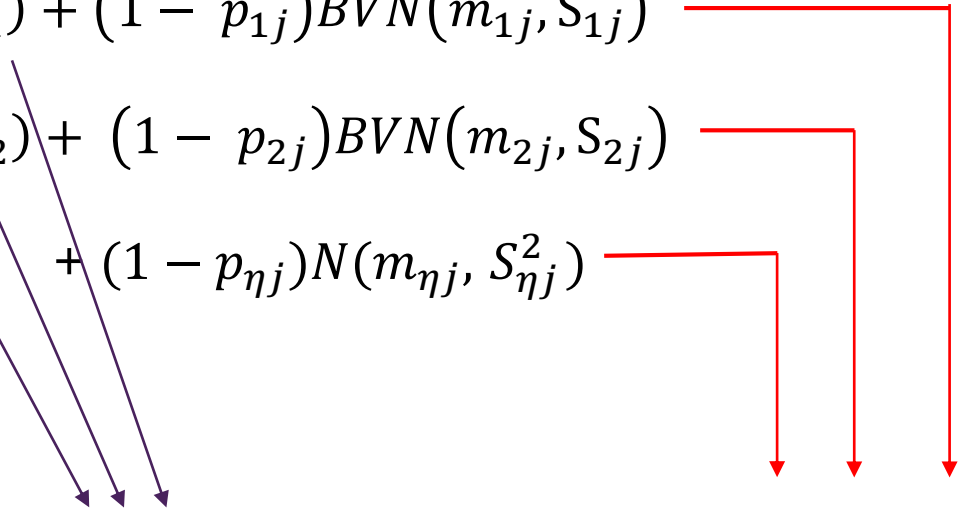
$$\text{odds}(\pi_{12,d_1,d_2}) = \text{odds}(\pi^0_{12,d_1,d_2})\exp(\eta d_1 d_2)$$

$$(\alpha_1, \beta_1, \alpha_2, \beta_2 > 0)$$

- (note: reference/scaling doses dropped in formulas)
- if no dose-dependent interaction desired: simply use *exp(η)*
- typically $\eta > 0$, but not necessarily

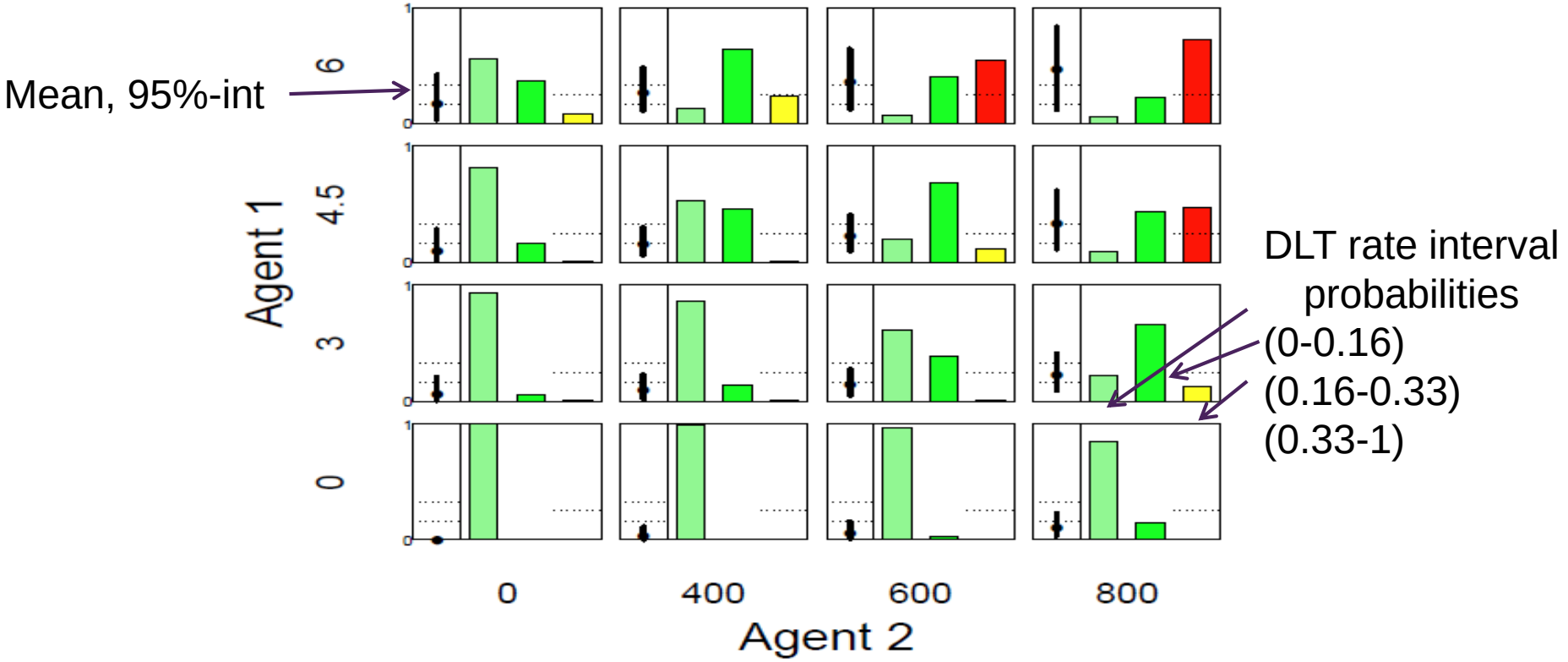
Robust Co-Data Model for Drug Combination Studies

- Let us assume $\theta_{1j} = (\log(\alpha_{1j}), \log(\beta_{1j}))$ and $\theta_{2j} = (\log(\alpha_{2j}), \log(\beta_{2j}))$
- $\theta_{1j} \sim p_{1j} BVN(\mu_1, \Gamma_1) + (1 - p_{1j}) BVN(m_{1j}, S_{1j})$
- $\theta_{2j} \sim p_{2j} BVN(\mu_2, \Gamma_2) + (1 - p_{2j}) BVN(m_{2j}, S_{2j})$
- $\eta_j \sim p_{\eta j} N(\mu_\eta, \tau_\eta^2) + (1 - p_{\eta j}) N(m_{\eta j}, S_{\eta j}^2)$



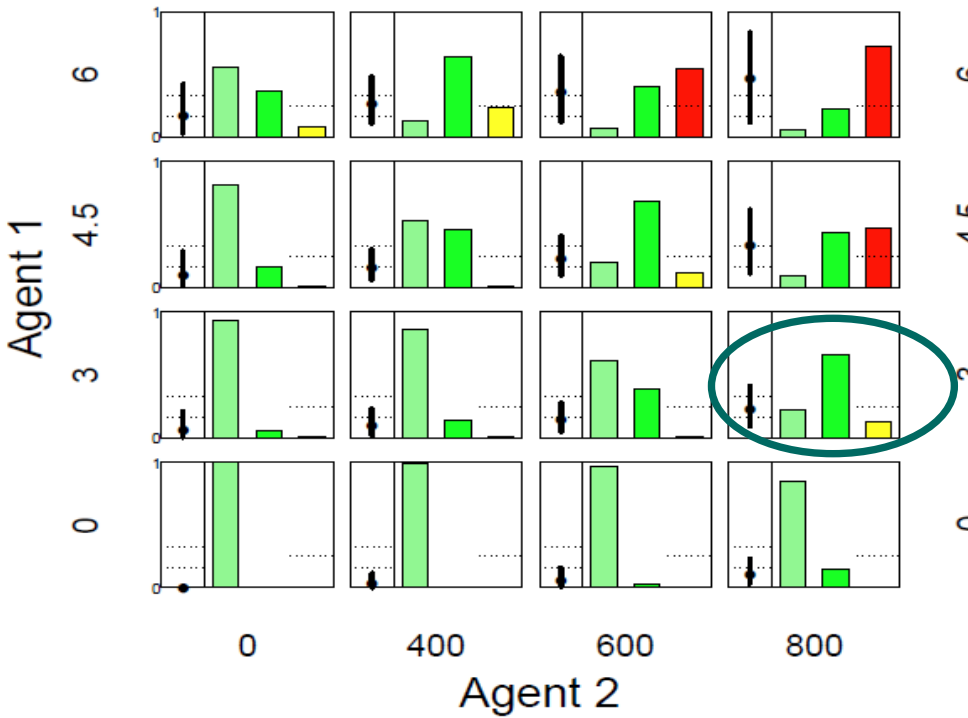
Exchangeable part **Non-exchangeable part**

Risk-Benefit Plot

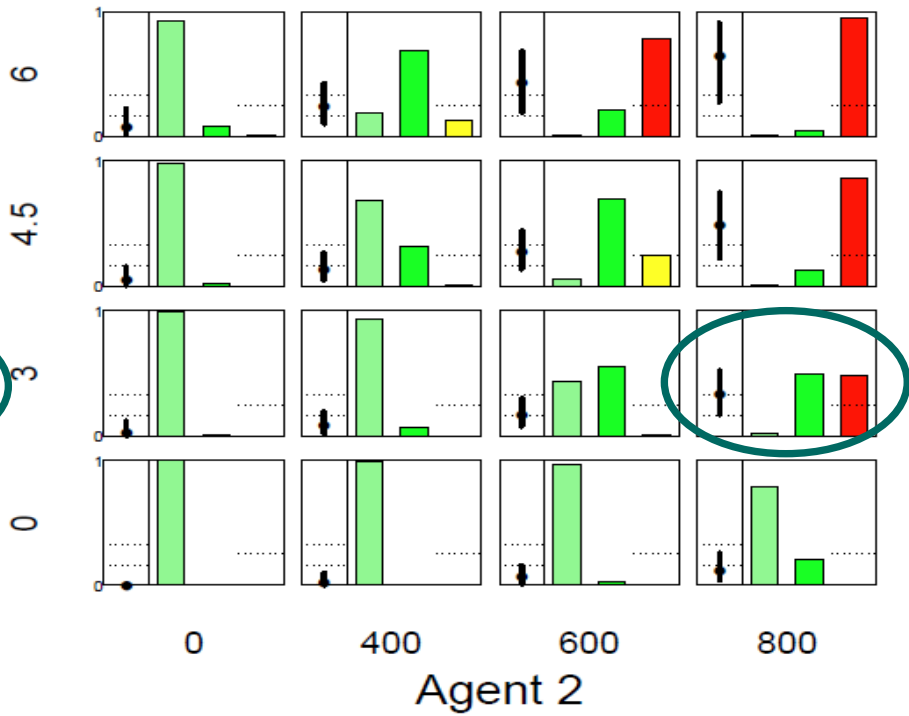


Co-data Analysis

Analysis with historical co-data only



Analysis with all the co-data



Effective sample sizes

Information gain from co-data: effective sample sizes (ESS)

dose combination	n	1) without *co-data*	2) with historical *co-data*	3) with all *co-data*
			ESS after 3 cohorts of AB trial	
3-800	3	7	11	11
4.5-600	3	7	9	9
6-400	3	6	8	8
			ESS at end of AB trial	
3-800	6	16	23	26
4.5-600	10	26	26	31
6-400	10	19	20	24

Use of Co-data: Planned vs Unplanned

- Clear specification of statistical analysis method before trial begins
- Proper choice of evidence is necessary
  - prior to start of trial
  - choice must be “science” based not “result” based
  - avoiding publication bias
  - inter-disciplinary collaboration

Conclusion

- Making better use of data - which includes *co-data* - is one contribution to innovation in medical product development
- Many applications with *co-data*
  - pediatric trials (adult data), non-inferiority trials (placebo, active control data), health-technology assessments, basket trials
- Methodology (meta-analytic) fairly well developed
- *Co-data* use: mainly for early phase trials or trial adaptations
  - what about using *co-data* for primary analysis in confirmatory trials?
  - not commonly used, but the mindset changes...
  - recent example in epilepsy (historical controls) Katz (2006), French (2010), Wechsler (2014)

Thank You