A Bayesian model averaging approach for basket trials

Junichi (Jun) Asano[1] and Akihiro Hirakawa[2]

[1] Biostatistics Group, Center for Product Evaluation, Pharmaceuticals and Medical Devices Agency (PMDA)

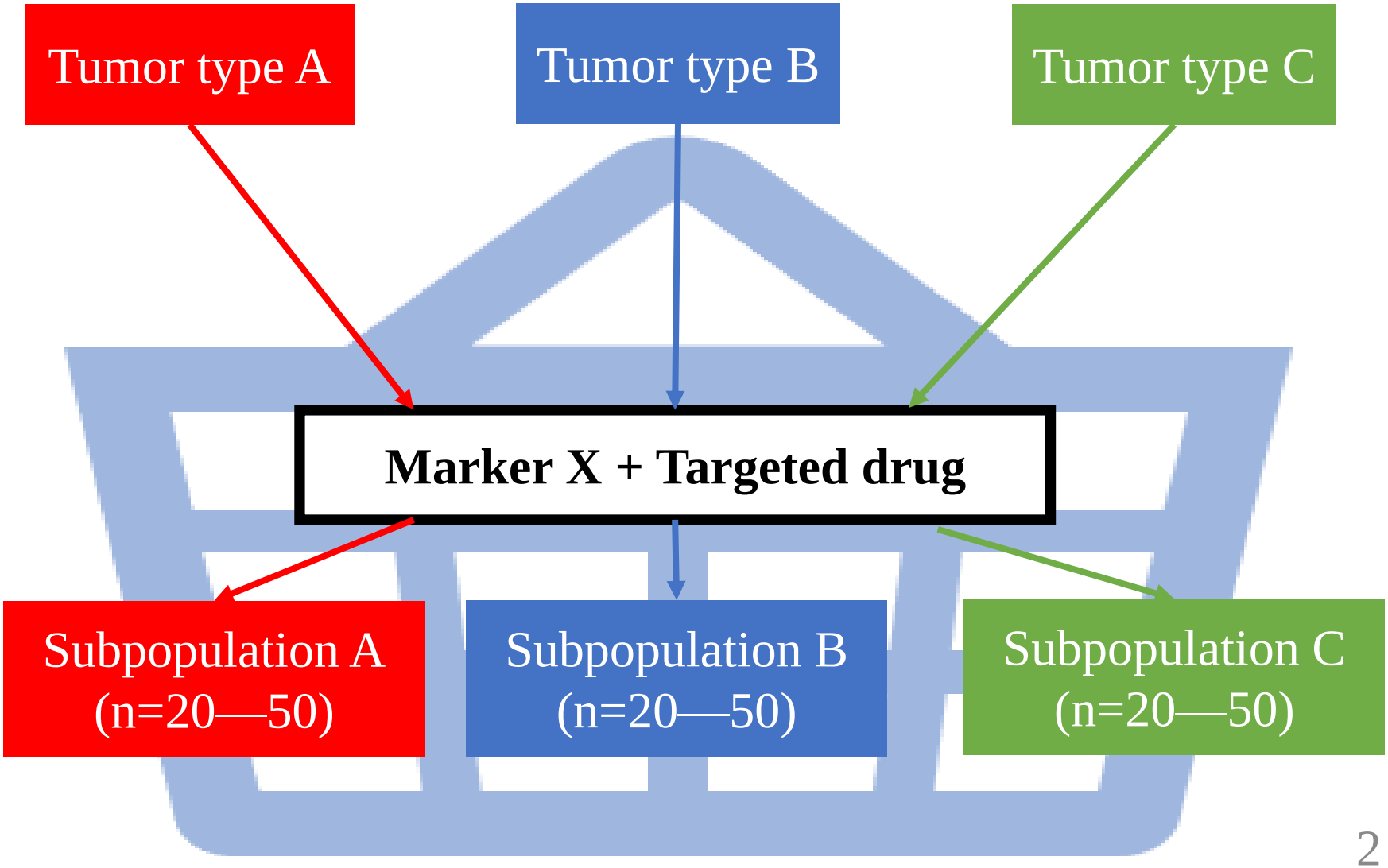
[2] Department of Biostatistics and Bioinformatics, Graduate School of Medicine, The University of Tokyo

Disclaimer

The views expressed in this presentation are those of the presenter and do not necessarily reflect the official views of Pharmaceuticals and Medical Devices Agency (PMDA)

The presenter have no conflicts of interest to declare

Single-arm basket trial in Phase II



Evaluation of response rate

- The response rate is often evaluated based on
  - Pooled analysis that estimates a response rate from all enrolled patients
  - Independent analysis that estimates a response rate for each cancer type

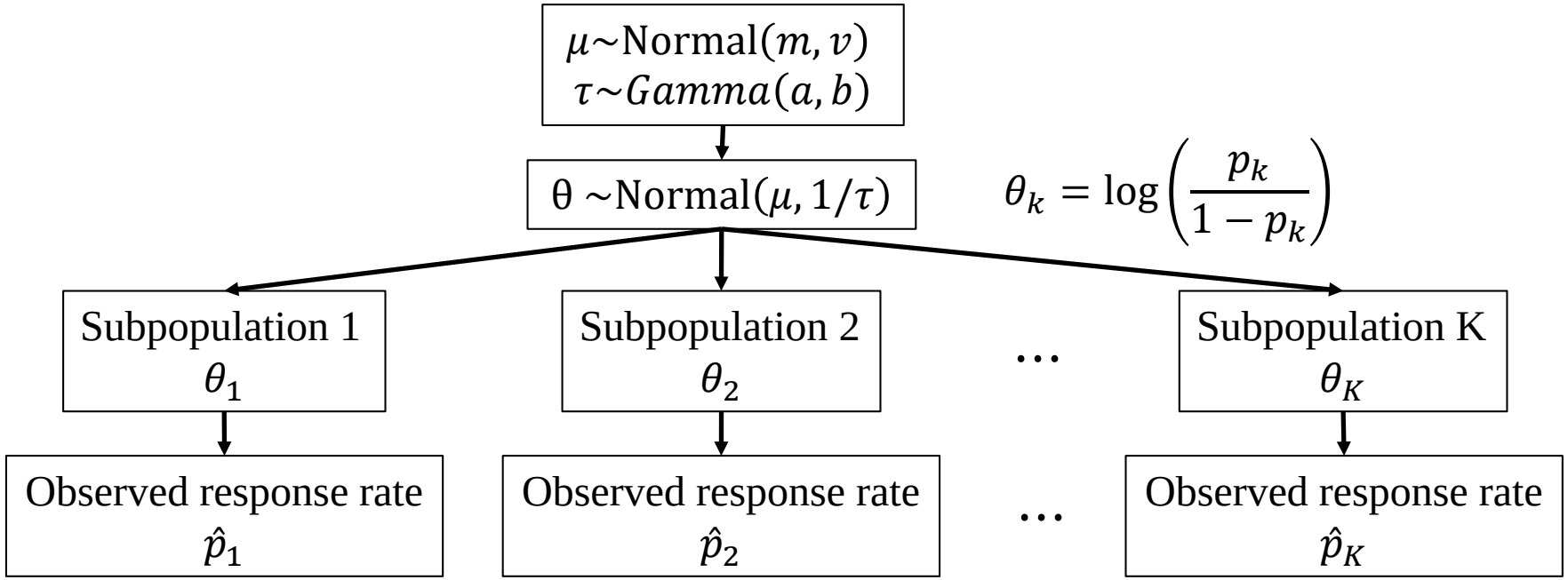
Subpopulation	k=1	2	…	K	Total
# of patients	n_1	n_2	…	n_K	n_{all}
# of responses	r_1	r_2	…	r_K	r_{all}
Observed response rate	$\hat{p}_1 = \frac{r_1}{n_1}$	$\hat{p}_2 = \frac{r_2}{n_2}$	…	$\hat{p}_K = \frac{r_K}{n_K}$	$\hat{p}_{all} = \frac{r_{all}}{n_{all}}$

Independent analysis

Pooled analysis

Bayesian hierarchical model

- Bayesian hierarchical model (BHM) that can borrow information related to treatment effects across cancer types
  - Thall et al. (2003), Berry et al., (2013)

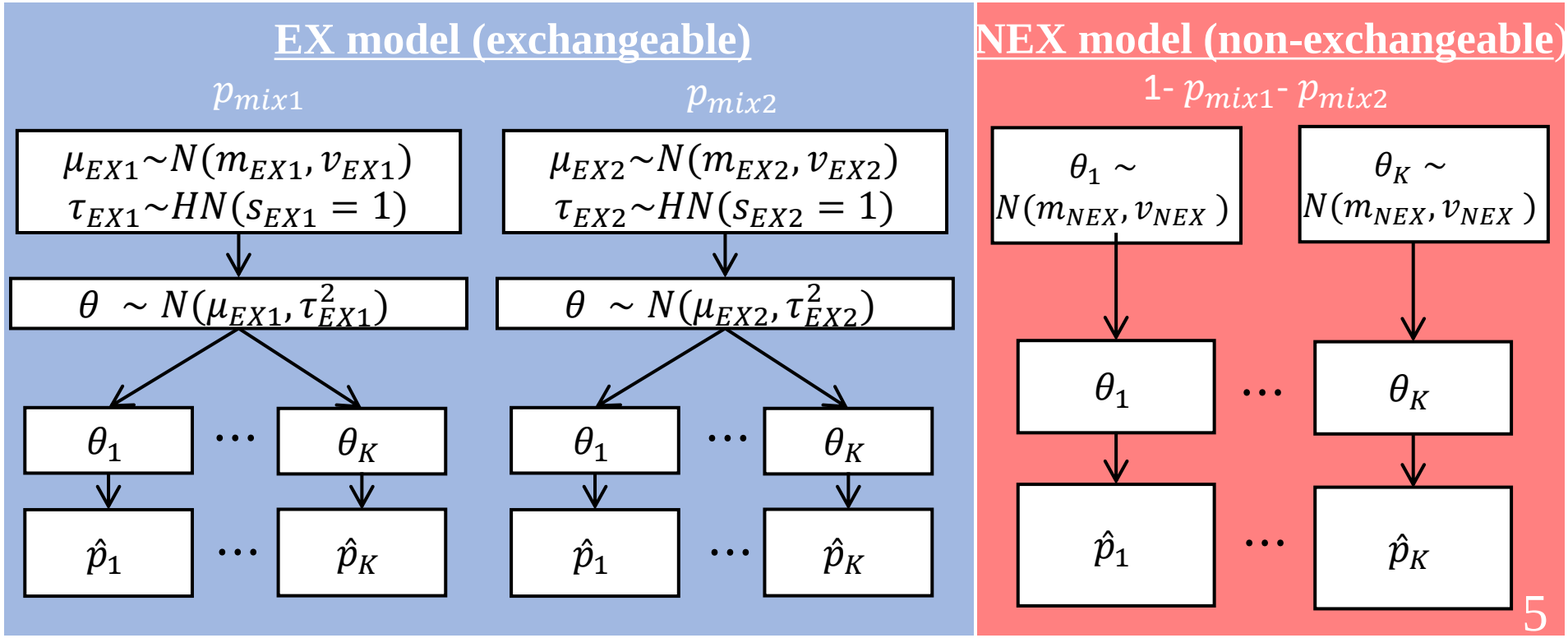


- Posterior distribution

$$f(\boldsymbol{\theta}, \mu, \tau | \boldsymbol{n}, \boldsymbol{r}) = \left\{ \prod_{k=1}^{K} L(\theta_k | n_k, r_k)\ f(\theta_k | \mu, \tau) \right\} f(\mu | \mathrm{m}, v) f(\tau | \mathrm{a}, b)$$

Exchangeability- non-exchangeability model

- Exchangeability- non-exchangeability (EXNEX) model that divides subpopulations as exchangeable with other similar subpopulations and non-exchangeable with subpopulations
  - Neuenschwander et al., 2016



Simon's concept of basket design (2018)

- True response rate in subpopulation k, p_k, is defined as a dichotomic parameter where it is either π_1 (i.e., null response probability) or π_2 (i.e., alternative response probability)
- Posterior probability $\Pr[p_k = \pi_2 | \mathbf{r}, \mathbf{n}]$ is evaluated under two different hypotheses according to the homogeneity (H_0) / heterogeneity (H_1) of true response rate, based on the law of total probability as follow:

$$\Pr[p_k = \pi_2 | \mathbf{r}, \mathbf{n}] = \Pr[p_k = \pi_2 | \mathbf{r}, \mathbf{n}, H_0]\Pr[H_0 | \mathbf{r}, \mathbf{n}] + \Pr[p_k = \pi_2 | \mathbf{r}, \mathbf{n}, H_1]\Pr[H_1 | \mathbf{r}, \mathbf{n}]$$

$H_0: p_1 = p_2 = \cdots = p_K = p$

Sub population	True response rate
1	p
2	p
3	p

$H_1: Not\ H_0$

Sub population	True response rate			
	Case 1	Case 2	Case 3	Case 4
1	p_1	$p_{1,2}$	$p_{1,3}$	p_1
2	p_2	$p_{1,2}$	p_2	$p_{2,3}$
3	p_3	p_3	$p_{1,3}$	$p_{2,3}$

- Before beginning of the trial, specifications of the following parameters is required:
  - Prior probability for H_0, $\Pr[H_0] = \lambda$
  - Prior probability of $\Pr[p_k = \pi_2] = \gamma$

Purpose

- Existing methods can be roughly categorized into two groups:
  1. Methods borrowing information of response rates among subpopulations by the BHM
  2. Methods accounting for the degrees of homogeneity of response rates among subpopulations under the null and alternative hypotheses of true response
- Although existing methods require specifications of the prior distribution on homogeneity of response rates among subpopulations before beginning the trial, it is difficult to establish practically

- In this study, we propose the method based on Bayesian model averaging (BMA) in order to account the degrees of homogeneity/heterogeneity of response rates among subpopulations
- Proposed method is based on the Simon's concept of basket design
- Proposed method does not require specifications of prior distribution on homogeneity of response rates among subpopulations before beginning the trial

Estimation of response rate

- Targeted posterior probability $\Pr[p_k \geq \phi|\mathbf{r},\mathbf{n}]$ is evaluated through two different hypotheses according to the homogeneity (H_0) / heterogeneity (H_1) of true response rate, based on the law of total probability as follows:

$$\Pr[p_k \geq \phi|\mathbf{r},\mathbf{n}] = \Pr[p_k \geq \phi|\mathbf{r},\mathbf{n},H_0]\Pr[H_0|\mathbf{r},\mathbf{n}] + \Pr[p_k \geq \phi|\mathbf{r},\mathbf{n},H_1]\Pr[H_1|\mathbf{r},\mathbf{n}]$$

  - H_0 is the hypothesis that true response rates in all subpopulations are equal
  - H_1 is not H_0, and under H_1 , we assume a two component mixture beta distribution for true response rate in each subpopulation

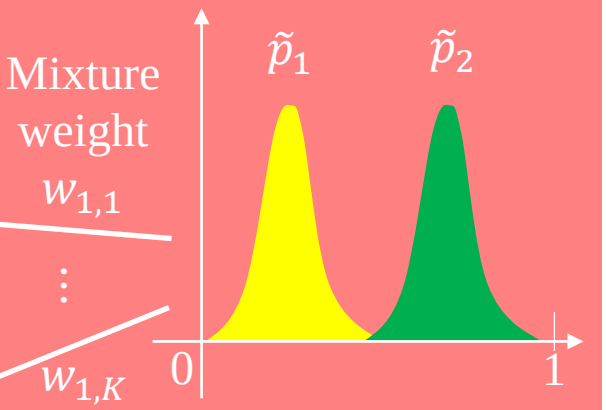
$H_0: p_1 = p_2 = \cdots = p_K = p_{all}$

Sub population	True response rate
1	p_{all}
…	…
K	p_{all}
Total	p_{all}

Pooled analysis

$H_1: Not\ H_0$

Sub population	True response rate
1	p_1
…	…
K	p_K



Two component mixture beta distribution of true response rate

Estimation of response rate

Targeted posterior probability

$$\Pr[p_k \geq \phi|\mathbf{r},\mathbf{n}] = \Pr[p_k \geq \phi|\mathbf{r},\mathbf{n},H_0]\Pr[H_0|\mathbf{r},\mathbf{n}] + \Pr[p_k \geq \phi|\mathbf{r},\mathbf{n},H_1]\Pr[H_1|\mathbf{r},\mathbf{n}]$$

- For the estimation of $\Pr[p_k \geq \phi|\mathbf{r},\mathbf{n},H_0]$, we use the following equation in common among subpopulations:

$$\Pr[p_k \geq \phi|\mathbf{r},\mathbf{n},H_0] = \int_{\phi}^{1} \binom{n_{all}}{r_{all}} p_{all}^{r_{all}}(1-p_{all})^{n_{all}-r_{all}} \frac{p_{all}^{\alpha-1}(1-p_{all})^{\beta-1}}{Beta(\alpha,\beta)} dp_{all}$$

  - Assume the binomial distribution with parameters n_{all} and p_{all} for r_{all}
  - Assume the prior beta distribution with hyperparameters α and β for p_{all}

- For the estimation of $\Pr[p_k \geq \phi|\mathbf{r},\mathbf{n},H_1]$, we use the following equation in each subpopulation:

$$\Pr[p_k \geq \phi|\mathbf{r},\mathbf{n},H_1] = \sum_{c=1}^{2} w_{c,k} \int_{\phi}^{1} \binom{n_k}{r_k} \tilde{p}_c^{r_k}(1-\tilde{p}_c)^{n_k-r_k} \frac{\tilde{p}_c^{\alpha_c-1}(1-\tilde{p}_c)^{\beta_c-1}}{Beta(\alpha_c,\beta_c)} d\tilde{p}_c$$

  - Assume the binomial distribution with parameters n_k and $\tilde{p}_c$ for r_k
  - Assume the two component mixture distribution with hyperparameters α_c and β_c for $\tilde{p}_c$, $(c = 1,2)$
  - $w_{1,k}$ and $w_{2,k}$ $(w_{1,k} + w_{2,k} = 1)$ represents the mixture weights of each distribution of true response rate

Estimation of response rate

Targeted posterior probability

$$\Pr[p_k \geq \phi|\mathbf{r},\mathbf{n}] = \Pr[p_k \geq \phi|\mathbf{r},\mathbf{n},H_0]\Pr[H_0|\mathbf{r},\mathbf{n}] + \Pr[p_k \geq \phi|\mathbf{r},\mathbf{n},H_1]\Pr[H_1|\mathbf{r},\mathbf{n}]$$

- Based on the Bayesian therorem, $\Pr[H_0|\mathbf{r},\mathbf{n}]$ can be obtained by

$$\Pr[H_0|\mathbf{r},\mathbf{n}] = \frac{Pr[\mathbf{r}|\mathbf{n},H_0]Pr[H_0]}{Pr[\mathbf{r}|\mathbf{n},H_0]Pr[H_0] + Pr[\mathbf{r}|\mathbf{n},H_1]Pr[H_1]}$$

- Place equal probability value for each hypothesis and use $\Pr[H_0] = \Pr[H_1] = 0.5$, because we assume there is no prior information available on the correct model
- For the estimate of $\Pr[r_k|n_k,H_0]$, we use the following equation in common among subpopulations:

$$\Pr[r_k|n_k,H_0] = \int_0^1 \binom{n_{all}}{r_{all}} p_{all}^{r_{all}}(1-p_{all})^{n_{all}-r_{all}} \frac{p_{all}^{\alpha-1}(1-p_{all})^{\beta-1}}{Beta(\alpha,\beta)} dp_{all}$$

- For the estimate of $\Pr[r_k|n_k,H_1]$, we use the following equation in each subpopulation *k*:

$$\Pr[r_k|n_k,H_1] = \sum_{c=1}^{2} w_{c,k} \int_0^1 \binom{n_k}{r_k} \tilde{p}_c^{r_k}(1-\tilde{p}_c)^{n_k-r_k} \frac{\tilde{p}_c^{\alpha_c-1}(1-\tilde{p}_c)^{\beta_c-1}}{Beta(\alpha_c,\beta_c)} d\tilde{p}_c$$

- $\Pr[H_1|\mathbf{r},\mathbf{n}]$ is $1 - \Pr[H_0|\mathbf{r},\mathbf{n}]$

Prior specification

- The proposed method requires specifications of the following prior values:
  - $Pr[H_0]\ (= 1 - Pr[H_1])$
  - $w_{1,k}\ (= 1 - w_{2,k})\ ,(k = 1, \cdots K)$
  - $\alpha, \beta,\ \alpha_c, \beta_c\ ,(c = 1,2)$

- Place equal probability value for each hypothesis and use $\Pr[H_0] = \Pr[H_1] = 0.5$, because we assume there is no prior information available on the correct model
- Set $w_{1,k}$ by using its expected posterior estimate
- Set $\alpha, \beta, \alpha_c, \beta_c$ by using empirical Bayesian theorem
- Therefore, the proposed method does not require specifications of prior distribution on homogeneity of response rates among subpopulations before beginning the trial

Prior specification

- Based on the results of our preliminary evaluation, we obtained the expected posterior estimate of $w_{1,k}$ according to the Bayesian paradigm.

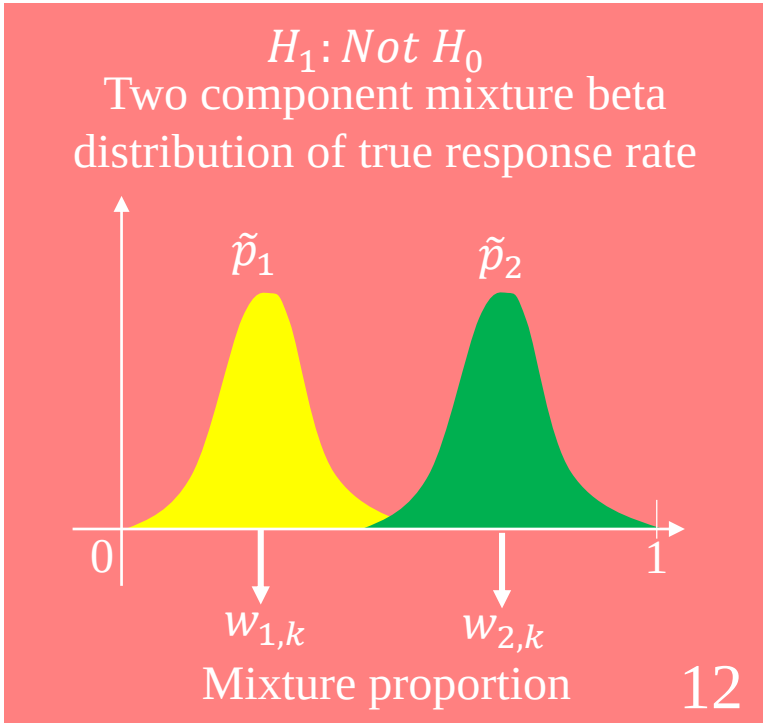
$$E(w_{1,k}) = \int_0^1 w_{1,k} \frac{f(r_k|n_k, w_{1,k}) f(w_{1,k})}{f(r_k|n_k, w_{1,k}) f(w_{1,k}) + f(r_k|n_k, w_{2,k}) f(w_{2,k})} dw_{1,k}$$

- Assume that $f(r_k|n_k, w_{1,k})$ and $f(r_k|n_k, w_{2,k})$ follow the binomial distribution with probability π_1 and π_2, respectively
- We also place equal probability of $f(w_{1,k}) = f(w_{2,k}) = 0.5$ since we assume there is no prior information available on the correct model.
- $E(w_{1,k})$ can be rewritten as:

$$E(w_{1,k}) = \int_0^1 w_{1,k} \left(1 + \left(\frac{\pi_2}{\pi_1}\right)^{r_k} \left(\frac{1-\pi_2}{1-\pi_1}\right)^{n_k - r_k}\right)^{-1} dw_{1,k}$$

$$= \frac{1}{2}\left(1 + \left(\frac{\pi_2}{\pi_1}\right)^{r_k} \left(\frac{1-\pi_2}{1-\pi_1}\right)^{n_k - r_k}\right)^{-1} \equiv \overline{w}_{1,k}$$

- Use $\overline{w}_{1,k}$ as the prior value for $w_{1,k}$ $(= 1 - w_{2,k})$



Prior specification

- For the prior values of α and β, we use the maximum likelihood estimates (MLEs) of $\hat{\alpha}$ and $\hat{\beta}$ obtained by maximizing the likelihood of the marginal distribution for p_{all}:

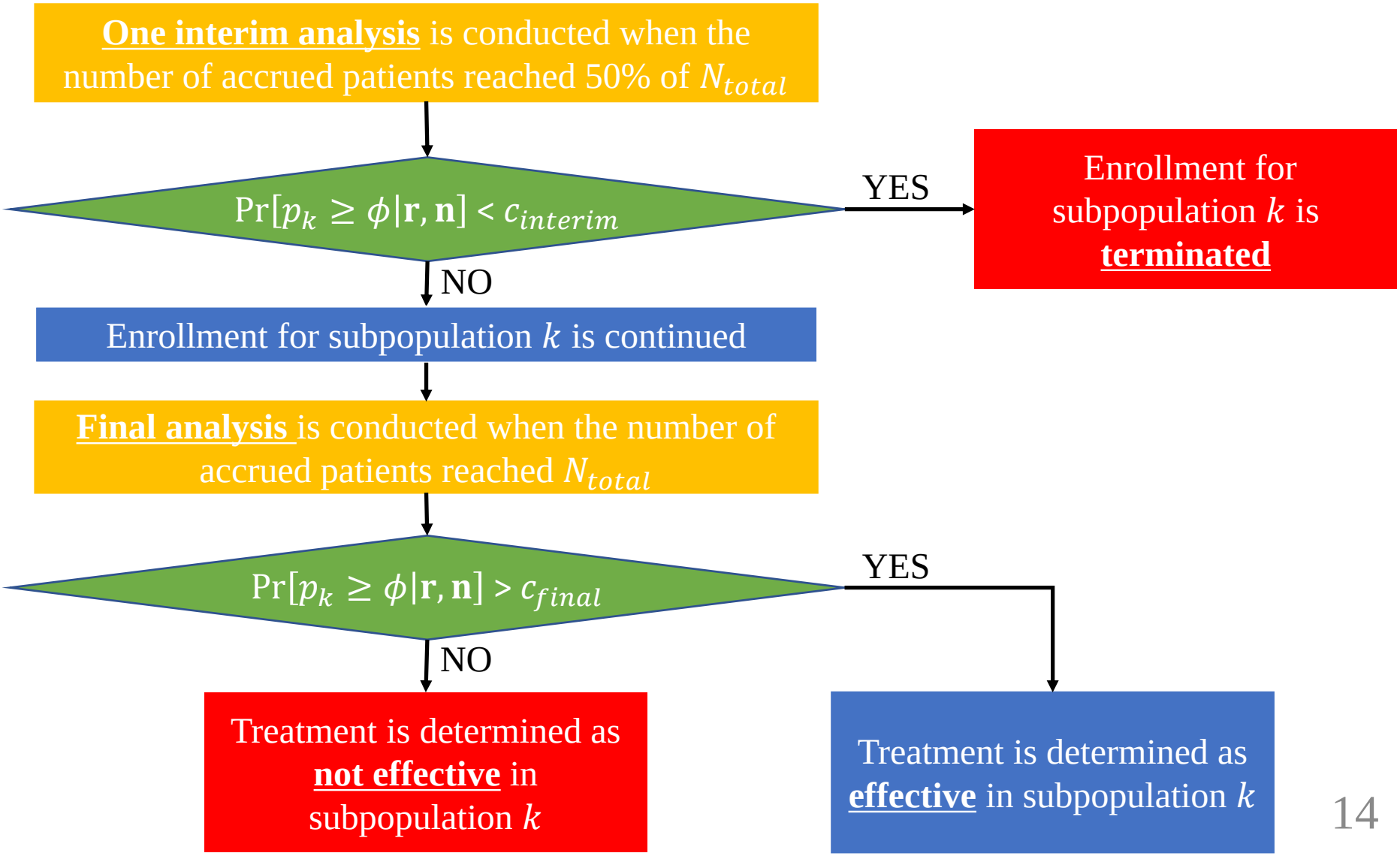
$$L[\alpha, \beta | \boldsymbol{r}, \mathbf{n}] = \int_0^1 \binom{n_{all}}{r_{all}} p_{all}^{r_{all}} (1 - p_{all})^{n_{all} - r_{all}} \frac{p_{all}^{\alpha - 1} (1 - p_{all})^{\beta - 1}}{Beta(\alpha, \beta)} \, dp_{all}$$

- For the prior values of α_c and β_c, (c = 1, 2) , we use the MLE of $\hat{\alpha}_c$ and $\hat{\beta}_c$ obtained by maximizing the likelihood of the marginal distribution for $\tilde{p}_c$:

$$L[\alpha_1, \beta_1, \alpha_2, \beta_2 | \boldsymbol{r}, \mathbf{n}, \bar{\boldsymbol{w}}_1]$$

$$= \prod_{k=1}^{K} \left\{ \bar{w}_{1,k} \int_0^1 \binom{n_k}{r_k} \tilde{p}_1^{r_k} (1 - p_1)^{n_k - r_k} \frac{\tilde{p}_1^{\alpha_1 - 1} (1 - \tilde{p}_1)^{\beta_1 - 1}}{Beta(\alpha_1, \beta_1)} \, d\tilde{p}_1 \right.$$

Trial design in simulation study



Simulation setting

- Null probability π_1 was 0.05
- Alternative probability π_2 was 0.25
- Targeted posterior probability was $\Pr[p_k \geq 0.05|\mathbf{r}, \mathbf{n}]$
- The number of subpopulations K was 4
- The total number of patients N_{total} was 100
- Futility stop criterion for the interim analysis $c_{interim}$ was 0.2

Simulation scenarios

Scenario	Subpopulation			
	1	2	3	4
1	**0.05**	**0.05**	**0.05**	**0.05**
2	0.25	**0.05**	**0.05**	**0.05**
3	0.25	0.25	**0.05**	**0.05**
4	0.25	0.25	0.25	**0.05**
5	0.25	0.25	0.25	0.25

- Efficiency criterion for the final analysis c_{final} was 0.9
- The maximum number of the patients enrolled into each subpopulation was identical at N_{upper} = 25
- $\mathbf{n} = (n_{1,\ldots,} n_K)^t$ was generated from a multinomial distribution with equal success probability
- Given the above-mentioned $\mathbf{n}$, $\mathbf{r} = (r_{1,\ldots,} r_K)^t$ was generated from a binomial distribution with true response probability
- We conducted 1,000 simulations for each scenario.

Models for comparison

- **Independent method**
  - Assumes the uninformative beta prior distribution which both parameter values were 0.5 for p_k
  - Obtain the posterior beta distribution of p_k based on the Bayesian theorem
- **Simon's method**
  - π_1 = 0.05, π_2 = 0.25
  - Prior probability for H_0, $\Pr[H_0] = \lambda$, was 0.33
  - Prior probability of $\Pr[p_k = \pi_2 | H_1] (= \gamma)$ was 0.5
- **BHM model**
  - $m = -1.734$ (i.e., $p_k = (\pi_1 + \pi_2)/2 = 0.15$) and $v = 10$
  - The weak degree of information borrowing : $(a, b) = (2, 20)$
  - The strong degree of information borrowing : $(a, b) = (2,2)$
- **EXNEX model**
  - $\mathrm{m}_{EX1} = -2.944$ (i.e., $p_k = 0.05$),$\mathrm{v}_{EX1} = 20.1$, and $s_{EX1} = 1$ for EX1
  - $\mathrm{m}_{EX2} = -1.098$ (i.e., $p_k = 0.25$),$\mathrm{v}_{EX2} = 4.3$, and $s_{EX2} = 1$ for EX2
  - $\mathrm{m}_{NEX} = -1.734$ (i.e.,$p_k = (\pi_1 + \pi_2)/2 = 0.15$) and $\mathrm{v}_{NEX} = 7.8$ for NEX
  - Mixture weights for the EXs and NEX components were equal

Simulation results

- Type I error rate and Power

Scenario	Method	Subpopulation 1	2	3	4
1	Proposed	5%	4%	4%	5%
	Independent	15%	11%	14%	11%
	Simon	0%	1%	1%	1%
	BHM-weak	5%	3%	4%	4%
	BHM-strong	5%	5%	5%	5%
	EXNEX	8%	6%	6%	6%
2	Proposed	88%	12%	11%	10%
	Independent	97%	11%	14%	11%
	Simon	79%	1%	1%	1%
	BHM-weak	90%	4%	4%	4%
	BHM-strong	91%	8%	11%	9%
	EXNEX	93%	9%	10%	9%
3	Proposed	95%	92%	12%	10%
	Independent	97%	94%	14%	11%
	Simon	79%	76%	2%	1%
	BHM-weak	91%	88%	4%	6%
	BHM-strong	94%	91%	17%	15%
	EXNEX	94%	92%	14%	11%
4	Proposed	96%	93%	94%	10%
	Independent	97%	95%	96%	11%
	Simon	80%	77%	78%	4%
	BHM-weak	91%	88%	89%	8%
	BHM-strong	97%	95%	96%	28%
	EXNEX	96%	94%	95%	11%
5	Proposed	95%	92%	93%	91%
	Independent	96%	94%	95%	94%
	Simon	87%	84%	85%	84%
	BHM-weak	92%	89%	90%	88%
	BHM-strong	99%	98%	99%	98%
	EXNEX	96%	94%	95%	94%

Simulation results

Scenario	Family wise error rate (FWER)					
	Method					
	Proposed	Independent	Simon	BHM-weak	BHM-strong	EXNEX
1	12%	42%	3%	15%	15%	19%
2	30%	32%	2%	12%	24%	24%
3	21%	24%	3%	10%	26%	23%
4	10%	11%	4%	8%	28%	11%
Average	18%	27%	3%	11%	23%	19%

Simulation results

- Futility stopping rate for interim analysis
  - Subpopulation with Null response probability
  - Subpopulation with Alternative response probability

Scenario	Method	Subpopulation 1	2	3	4
1	Proposed	51%	52%	50%	54%
	Independent	6%	6%	5%	8%
	Simon	64%	65%	63%	65%
	BHM-weak	51%	52%	50%	54%
	BHM-strong	31%	29%	29%	32%
	EXNEX	36%	35%	35%	39%
2	Proposed	3%	50%	48%	52%
	Independent	0%	6%	5%	8%
	Simon	5%	64%	61%	63%
	BHM-weak	3%	52%	50%	54%
	BHM-strong	2%	14%	14%	14%
	EXNEX	2%	29%	28%	32%
3	Proposed	3%	5%	48%	51%
	Independent	0%	0%	5%	8%
	Simon	5%	6%	56%	57%
	BHM-weak	3%	5%	50%	54%
	BHM-strong	1%	1%	3%	2%
	EXNEX	1%	2%	21%	23%
4	Proposed	3%	4%	4%	52%
	Independent	0%	0%	0%	8%
	Simon	4%	5%	5%	41%
	BHM-weak	3%	5%	4%	54%
	BHM-strong	0%	0%	0%	0%
	EXNEX	1%	1%	1%	15%
5	Proposed	3%	4%	4%	5%
	Independent	0%	0%	0%	0%
	Simon	3%	2%	3%	3%
	BHM-weak	3%	5%	4%	5%
	BHM-strong	0%	0%	0%	0%
	EXNEX	0%	0%	0%	0%

Conclusion

- The proposed method would be useful in the practical setting of “signal-finding” basket trials without prior information on the treatment effect of the candidate drug because it does not require specifications of prior distribution on homogeneity of response rates among subpopulations

- On an average, the proposed methods displayed an intermediate performance between the BHM-weak and BHM-strong

- In-depth understanding of the homogeneity of the treatment effect across subpopulations obtained using multiple approaches with different assumptions and/or methodological features is important

Thank you

Mt. Fuji, Japan

Mt. Hood, Oregon

References

- Thall PF, Wathen JK, Bekele BN et al. Hierarchical Bayesian approaches to phase II trials in diseases with multiple subtypes. *Statistics in Medicine* 2003; **22**: 763-780.
- Berry SM, Broglio KR, Groshen S, Berry DA. Bayesian hierarchical modeling of patient subpopulations: efficient designs of phase II oncology clinical trials. *Clinical Trials* 2013; **10**: 720-734.
- Neuenschwander B, Wandel S, Roychoudhury S, Bailey S. Robust exchangeability designs for early phase clinical trials with multiple strata. *Pharmaceutical Statistics* 2016; **15**: 123-134.
- Simon R, Geyer S, Subramanian J, Roychowdhury S. The Bayesian basket design for genomic variant-driven phase II trials. *Seminars in Oncology* 2016; **43**: 13-18.
- Simon R. New designs for basket clinical trials in oncology. *Journal of Biopharmaceutical Statistics* 2018; **28**: 245-255.

Back up

Simulation results

Method	Average Type I error rate	Average Power
Proposed	8%	93%
Independent	12%	95%
Simon	1%	81%
BHM-weak	5%	90%
BHM-strong	11%	96%
EXNEX	9%	94%

Scenario	Family wise error rate (FWER)					
	Method					
	Proposed	Independent	Simon	BHM-weak	BHM-strong	EXNEX
1	12%	42%	3%	15%	15%	19%
2	30%	32%	2%	12%	24%	24%
3	21%	24%	3%	10%	26%	23%
4	10%	11%	4%	8%	28%	11%
Average	18%	27%	3%	11%	23%	19%

Simulation results

Average futility stopping rate for interim analysis		
Method	Subpopulation with true response probability of 0.05 (Null)	Subpopulation with true response probability of 0.25 (Alternative)
Proposed	51%	4%
Independent	7%	0%
Simon	60%	4%
BHM-weak	52%	4%
BHM-strong	17%	0%
EXNEX	29%	1%

	Number of enrolled patients					
	Method					
Scenario	Proposed	Independent	Simon	BHM-weak	BHM-strong	EXNEX
1	73	93	68	73	81	80
2	79	93	76	78	89	86
3	84	93	83	84	93	90
4	88	93	89	88	93	92
5	92	93	92	92	93	93
Average	83	93	82	83	90	88

BRAF V600E trial

Subpopulation	Number of patients	Number of response	Observed response rate
ECD/LCH	14	6	43%
NSCLC	19	8	42%
Anaplastic Thyroid Cancer	7	2	29%
Cholangio-carcinoma	8	1	13%
Colorectal	10	0	0%
Total	58	17	29%

BRAF V600E trial

- According to the statistical settings of this trial, we set π_1, π_2, and ϕ to 0.15, 0.45, and 0.35, respectively
- Estimated the posterior probability $\Pr[p_k \geq 0.35|\mathbf{r}, \mathbf{n}]$ for each method
- m = -0.8473 for p_k = (0.15+0.45)/2 = 0.3 was commonly used in the BHM methods
- $m_{EX1} = -1.735$ (i.e., $p_k = 0.15$), $m_{EX2} = -0.201$ (i.e., $p_k = 0.45$),and $m_{NEX} = -0.847$ (i.e., $p_k = 0.3$) in the EXNEX method
- the values for the prior probability γ was set to 0.1 or 0.5 in the Simon's method.
- The other configurations for each method were the same as those used in the simulation studies

BRAF V600E trial

Subpopulation	Observed	Proposed	Independent	Simon (γ=0.5)	Simon (γ=0.1)	BHM-weak	BHM-strong	EXNEX
ECD/LCH	43% (6/14)	61%	74%	95%	69%	70%	55%	68%
NSCLC	42% (8/19)	59%	75%	98%	83%	72%	60%	68%
Anaplastic Thyroid Cancer	29% (2/7)	31%	38%	61%	20%	31%	25%	35%
Cholangio-carcinoma	13% (1/8)	7%	8%	31%	11%	5%	8%	9%
Colorectal	0% (0/10)	1%	0%	13%	8%	0%	1%	1%
Total	29% (17/58)							

Simulation scenarios (K=8)

Scenario	Subpopulation							
	1	2	3	4	5	6	7	8
1	**0.05**	**0.05**	**0.05**	**0.05**	**0.05**	**0.05**	**0.05**	**0.05**
2	0.25	**0.05**	**0.05**	**0.05**	**0.05**	**0.05**	**0.05**	**0.05**
3	0.25	0.25	**0.05**	**0.05**	**0.05**	**0.05**	**0.05**	**0.05**
4	0.25	0.25	0.25	**0.05**	**0.05**	**0.05**	**0.05**	**0.05**
5	0.25	0.25	0.25	0.25	**0.05**	**0.05**	**0.05**	**0.05**
6	0.25	0.25	0.25	0.25	0.25	**0.05**	**0.05**	**0.05**
7	0.25	0.25	0.25	0.25	0.25	0.25	**0.05**	**0.05**
8	0.25	0.25	0.25	0.25	0.25	0.25	0.25	**0.05**
9	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25

Simulation results (K=8)

Scenario	Method	Subpopulation							
		1	2	3	4	5	6	7	8
1	Proposed	2	4	3	2	3	2	3	3
	Independent	11	12	12	11	11	11	11	12
	Simon	1	1	1	1	1	1	2	1
	BHM-weak	3	4	3	3	4	3	4	3
	BHM-strong	3	2	3	3	2	2	2	2
	EXNEX	5	6	5	4	5	4	4	4
2	Proposed	70	7	6	5	5	5	5	5
	Independent	83	12	12	11	11	11	11	12
	Simon	64	1	1	1	2	1	2	1
	BHM-weak	75	4	4	3	5	3	4	3
	BHM-strong	61	5	5	5	5	5	4	4
	EXNEX	73	6	7	6	6	5	5	6
3	Proposed	75	73	6	6	6	6	6	5
	Independent	83	82	12	11	11	11	11	12
	Simon	63	61	1	1	2	1	1	1
	BHM-weak	74	72	4	4	5	4	5	3
	BHM-strong	69	70	8	8	8	8	6	7
	EXNEX	75	77	9	8	8	7	6	7

Simulation results (K=8)

Scenario	Method	Subpopulation							
		1	2	3	4	5	6	7	8
4	Proposed	77	75	77	6	7	5	7	6
	Independent	83	82	85	11	11	11	11	12
	Simon	63	61	62	1	2	1	1	1
	BHM-weak	74	73	75	3	4	4	5	3
	BHM-strong	77	76	78	11	12	12	10	10
	EXNEX	78	78	80	9	9	10	8	9
5	Proposed	77	75	76	76	8	7	7	6
	Independent	83	82	85	82	11	11	11	12
	Simon	59	60	63	60	2	1	2	1
	BHM-weak	74	72	74	74	6	6	5	4
	BHM-strong	83	81	83	80	16	17	17	17
	EXNEX	80	81	82	80	10	10	9	11
6	Proposed	77	76	78	76	76	7	6	8
	Independent	83	82	85	82	81	12	11	12
	Simon	60	60	63	60	61	2	3	2
	BHM-weak	74	73	75	75	73	6	5	5
	BHM-strong	86	85	88	84	84	25	23	24
	EXNEX	82	82	84	81	80	12	11	13

Simulation results (K=8)

Scenario	Method	Subpopulation							
		1	2	3	4	5	6	7	8
7	Proposed	76	77	79	77	75	75	7	8
	Independent	83	82	85	82	81	80	11	12
	Simon	62	63	64	63	61	62	5	5
	BHM-weak	75	76	76	75	74	71	6	5
	BHM-strong	90	89	91	90	88	89	34	36
	EXNEX	84	83	86	82	81	81	13	14
8	Proposed	76	78	80	77	75	76	74	8
	Independent	83	82	85	82	81	80	79	12
	Simon	70	70	70	69	68	68	66	14
	BHM-weak	75	78	78	76	74	72	74	6
	BHM-strong	94	93	94	94	93	93	92	50
	EXNEX	85	84	86	83	82	81	81	15
9	Proposed	78	77	79	76	76	76	74	77
	Independent	83	82	85	82	81	80	79	82
	Simon	82	81	83	79	79	79	78	79
	BHM-weak	77	76	79	76	76	75	74	77
	BHM-strong	96	96	97	97	95	96	95	96
	EXNEX	86	86	88	85	84	83	83	84

Simulation results (K=8)

Method	Average Type I error rate	Average Power
Proposed	5	76
Independent	11	82
Simon	2	67
BHM-weak	4	75
BHM-strong	11	87
EXNEX	8	82

Simulation results (K=8)

	FWERs					
	Method					
Scenario	Proposed	Independent	Simon	BHM-weak	BHM-strong	EXNEX
1	15	62	9	22	13	28
2	32	57	8	22	22	30
3	31	51	8	21	30	33
4	27	46	7	18	37	35
5	25	39	6	19	43	34
6	20	31	6	15	45	31
7	14	21	9	10	53	25
8	8	12	14	6	50	15
Average	21	40	8	17	36	29

Simulation results (K=8)

Scenario	Method	Subpopulation							
		1	2	3	4	5	6	7	8
1	Proposed	72	72	71	74	72	73	75	73
	Independent	0	0	0	0	0	0	0	0
	Simon	43	38	41	43	40	40	40	40
	BHM-weak	70	70	70	73	70	71	73	71
	BHM-strong	30	29	29	31	30	30	30	31
	EXNEX	20	20	20	21	20	21	20	21
2	Proposed	18	70	70	73	70	71	72	70
	Independent	0	0	0	0	0	0	0	0
	Simon	8	38	40	42	39	40	39	39
	BHM-weak	18	69	68	72	69	70	71	69
	BHM-strong	8	14	15	15	16	16	15	16
	EXNEX	4	13	14	14	14	14	13	14
3	Proposed	18	19	66	69	66	67	67	66
	Independent	0	0	0	0	0	0	0	0
	Simon	8	7	38	41	37	38	37	37
	BHM-weak	17	18	65	68	66	67	68	65
	BHM-strong	4	4	8	7	7	7	7	7
	EXNEX	2	2	9	9	9	9	8	11

Simulation results (K=8)

Scenario	Method	Subpopulation							
		1	2	3	4	5	6	7	8
4	Proposed	15	18	17	64	61	62	62	61
	Independent	0	0	0	0	0	0	0	0
	Simon	8	6	6	37	33	34	33	33
	BHM-weak	16	18	16	66	61	63	63	61
	BHM-strong	1	1	1	2	2	2	2	2
	EXNEX	2	1	1	5	5	4	4	4
5	Proposed	14	16	14	15	55	56	55	53
	Independent	0	0	0	0	0	0	0	0
	Simon	7	6	5	5	27	28	27	28
	BHM-weak	14	16	14	15	54	56	55	54
	BHM-strong	0	0	0	1	1	1	1	1
	EXNEX	1	0	1	1	2	3	2	3
6	Proposed	13	12	11	12	13	50	49	47
	Independent	0	0	0	0	0	0	0	0
	Simon	5	5	4	4	5	22	21	21
	BHM-weak	13	13	12	12	13	48	48	46
	BHM-strong	0	0	0	0	0	0	0	0
	EXNEX	0	0	1	0	0	1	2	2

Simulation results (K=8)

Scenario	Method	Subpopulation							
		1	2	3	4	5	6	7	8
7	Proposed	11	11	9	10	11	11	42	42
	Independent	0	0	0	0	0	0	0	0
	Simon	4	4	3	3	3	3	14	14
	BHM-weak	11	10	9	9	10	11	37	36
	BHM-strong	0	0	0	0	0	0	0	0
	EXNEX	0	0	0	0	0	0	1	1
8	Proposed	9	10	8	9	9	9	9	37
	Independent	0	0	0	0	0	0	0	0
	Simon	2	2	2	1	2	2	2	8
	BHM-weak	7	7	6	6	6	8	8	26
	BHM-strong	0	0	0	0	0	0	0	0
	EXNEX	0	0	0	0	0	0	0	1
9	Proposed	8	8	7	6	7	8	8	7
	Independent	0	0	0	0	0	0	0	0
	Simon	1	1	2	1	1	1	1	1
	BHM-weak	5	5	4	4	5	4	5	4
	BHM-strong	0	0	0	0	0	0	0	0
	EXNEX	0	0	0	0	0	0	0	0

Simulation results (K=8)

Method	Subpopulation with true response probability of 0.05 (Null)	Subpopulation with true response probability of 0.25 (Alternative)
Proposed	63	11
Independent	0	0
Simon	34	4
BHM-weak	62	10
BHM-strong	11	1
EXNEX	10	0

Simulation results (K=8)

	Method					
Scenario	Proposed	Independent	Simon	BHM-weak	BHM-strong	EXNEX
1	83	100	100	84	89	97
2	90	100	100	90	96	99
3	94	100	100	94	98	99
4	97	100	100	97	100	99
5	99	100	100	99	100	99
6	100	100	100	99	100	99
7	100	100	100	100	100	99
8	100	100	100	100	100	99
9	100	100	100	100	100	99
Average	96	100	100	96	98	99