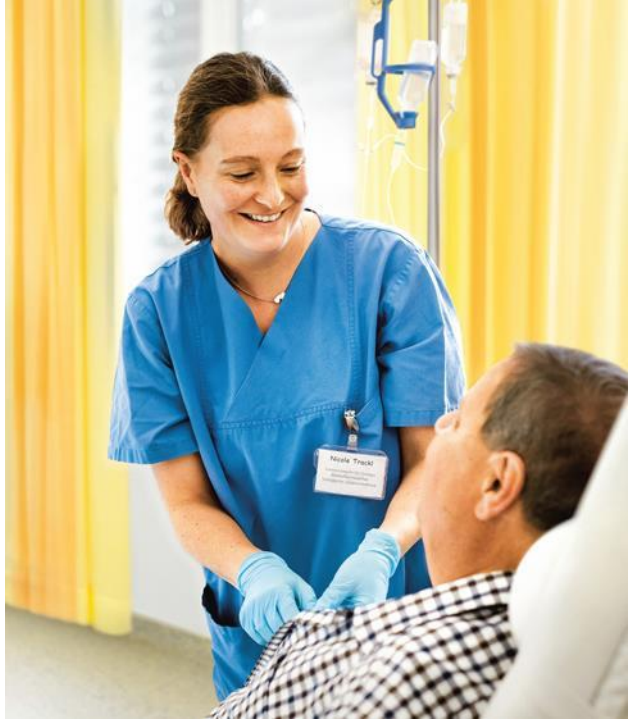




Science For A Better Life

Increasing Efficiency of Oncology Basket Trials using Bayesian Approach

Rong Liu
Pharmaceuticals Statistics, Bayer HealthCare Pharmaceuticals Inc., Whippany NJ

4th ANNUAL ASA NJ CHAPTER/BAYER STATISTICS WORKSHOP
11/11/2016

Acknowledgment

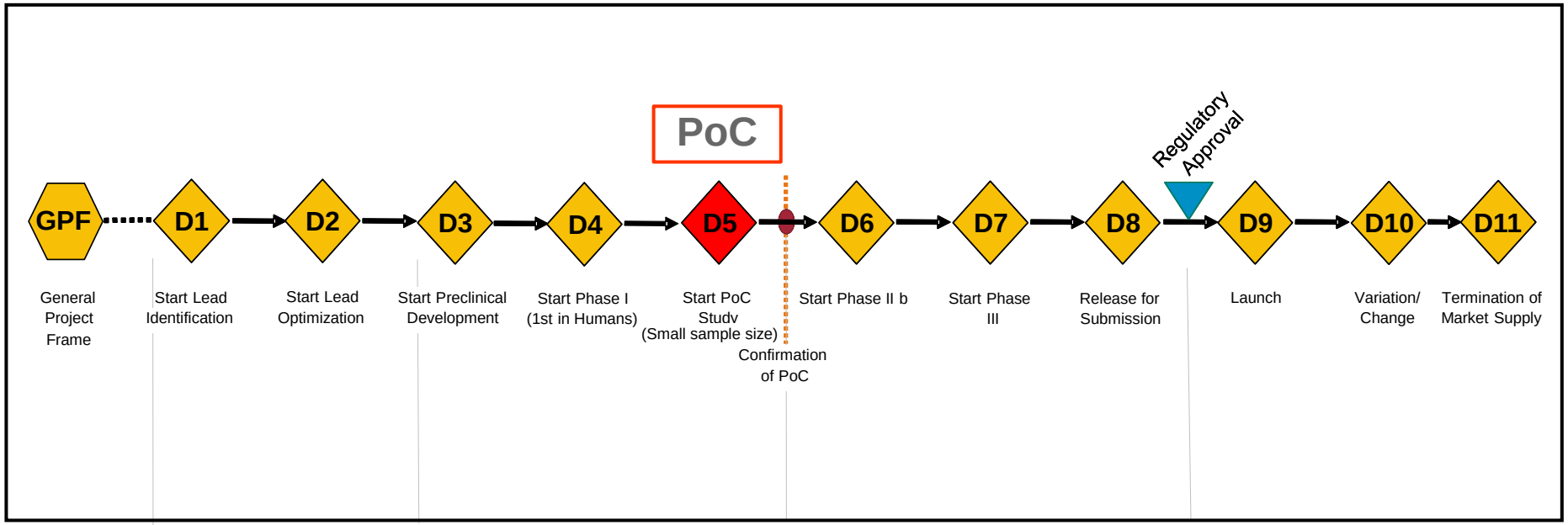
- Co-authors on the present work:
  - Alex Liu (UT Health Science Center at Houston)
  - Mercedeh Ghadessi (Bayer)
  - Richard Vonk (Bayer)

- We appreciate the encouragement and support from Sylvia Engelen and Daniel Haverstock from Bayer on this work

Agenda

- Introduction
  - Cancer Drug Development
  - Basket Trial Design
  - Bayesian Hierarchical Modeling
- Proposed Novel Design for Basket Trial using Bayesian Hierarchical Mixture Modeling
- Simulation Studies
- Summary

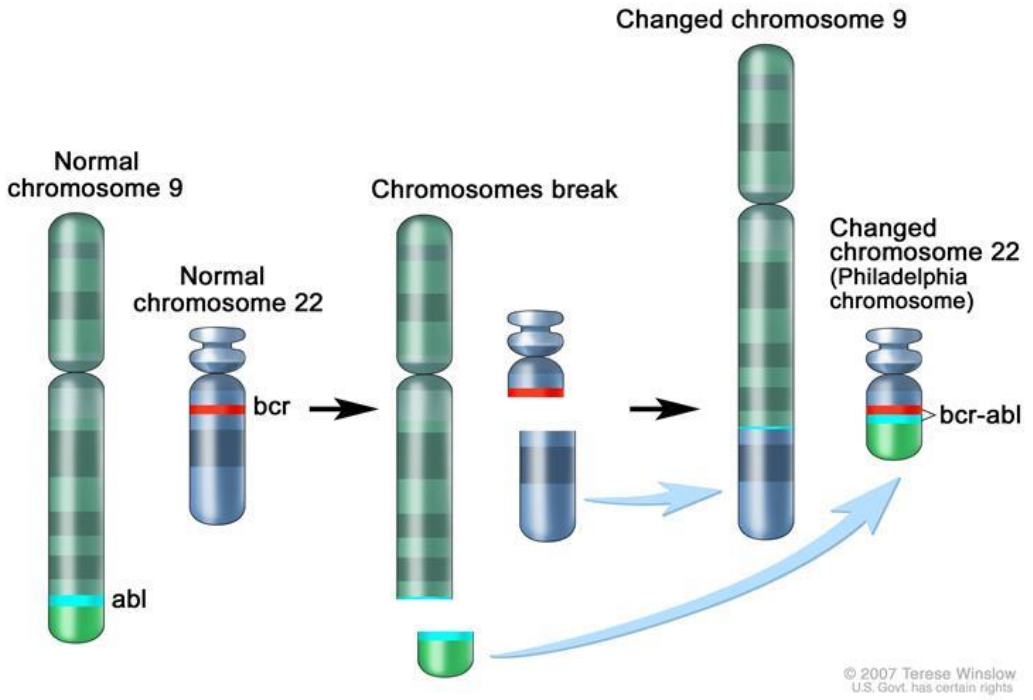
Cancer Drug Development Overview



- For Proof of Concept (POC) study, usually single arm small-scale studies to detect efficacy signal and are evaluated based on clinical and imaging criteria such as response rate (RR).
- Cancer is a disease that has been characterized and investigated separately based on the anatomic location: **More than 200 different types of cancer are determined based on the anatomic location![1]**

Targeted Therapy in Oncology

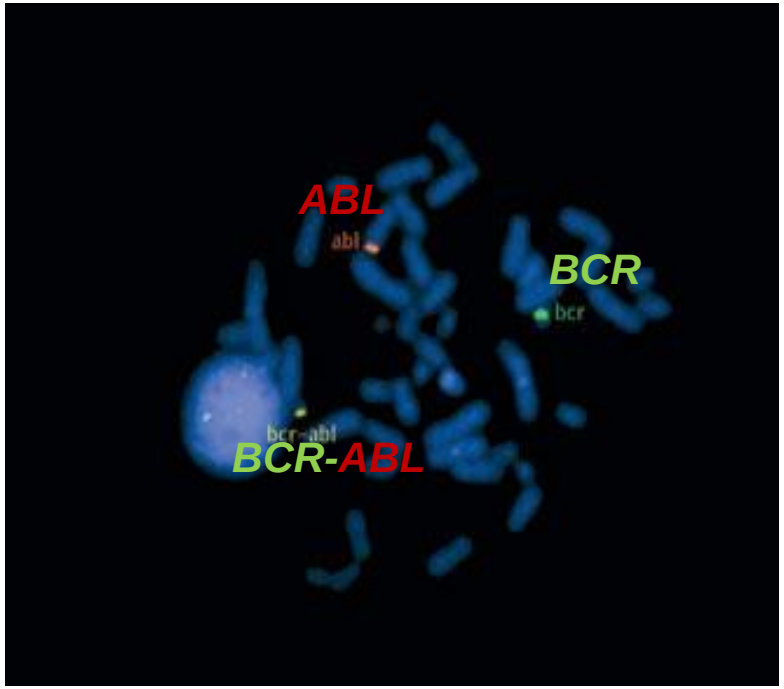
- In last decade, researcher have realized that majority of cancers have genetic risk factors[2]
  - BCR-ABL translocation, two chromosomes switch places (9 and 22)
  - Results in a “fusion gene” created by dis-positioning on ABL and BCR (**BCR-ABL Cancers)**



Targeted Therapy in Oncology

- **BCR-ABL Cancers** can be found in multiple cancer types
  - Chronic myeloid leukemia (CML)
  - Gastrointestinal stromal tumor (GIST)
  - Acute lymphoblastic leukemia (ALL)
  - Acute myelogenous leukemia (AML)

- Conducted clinical trial studies separately for CML, GIST, ALL, and AML[3-5]



http://path.svhm.org.au/services/Pages/Cytogenetics.aspx

Basket Trial

American Association for Cancer Research Cancer Progress Report 2015

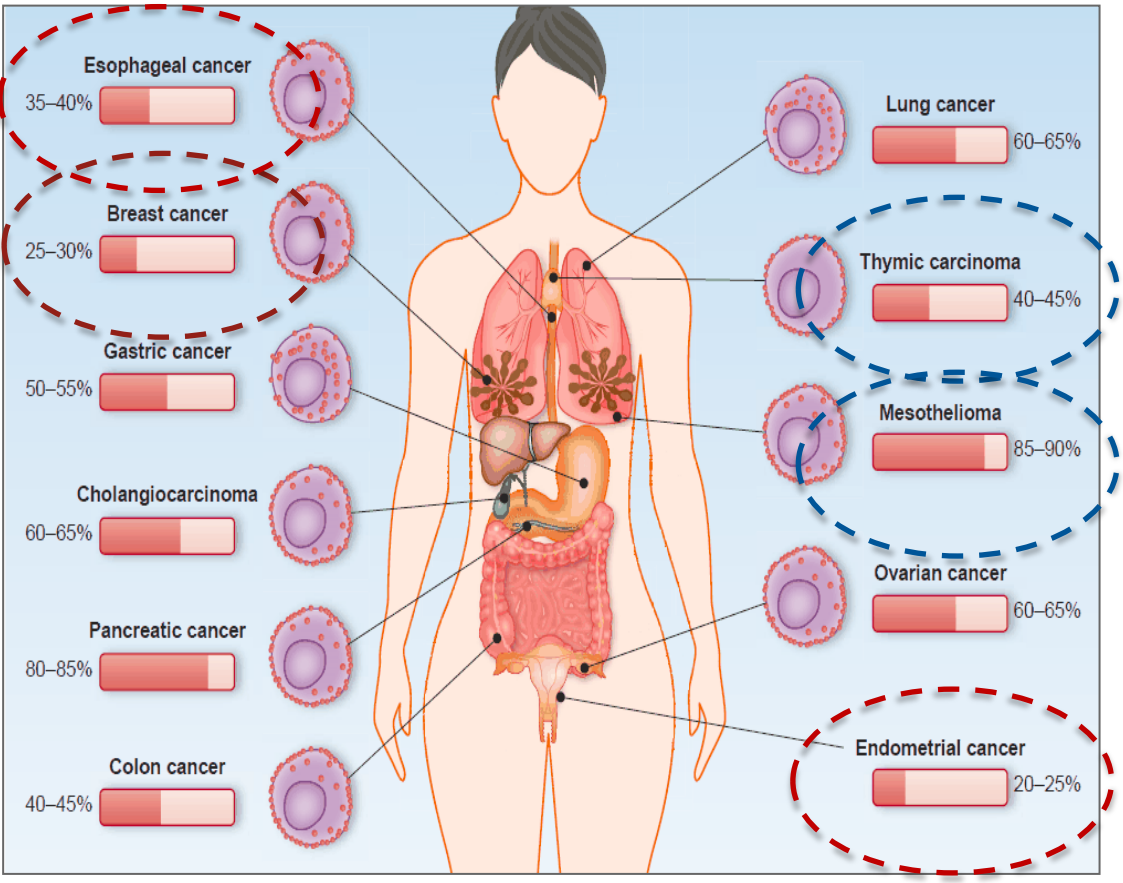
What is basket trial[7]?

- Trials based on genomics as opposed to site of origin
- Combing multiple cancer types in a single trial
- Molecular biomarker-selected and molecular subtype is more fundamental than histology
- Identify favorable response with a small number of patients

Answer the questions:

1) Does the treatment work on all studied cancer types?
2) If no, can we identify any cancer types with promising effect?

Role of Basket Trials in Targeted Therapy

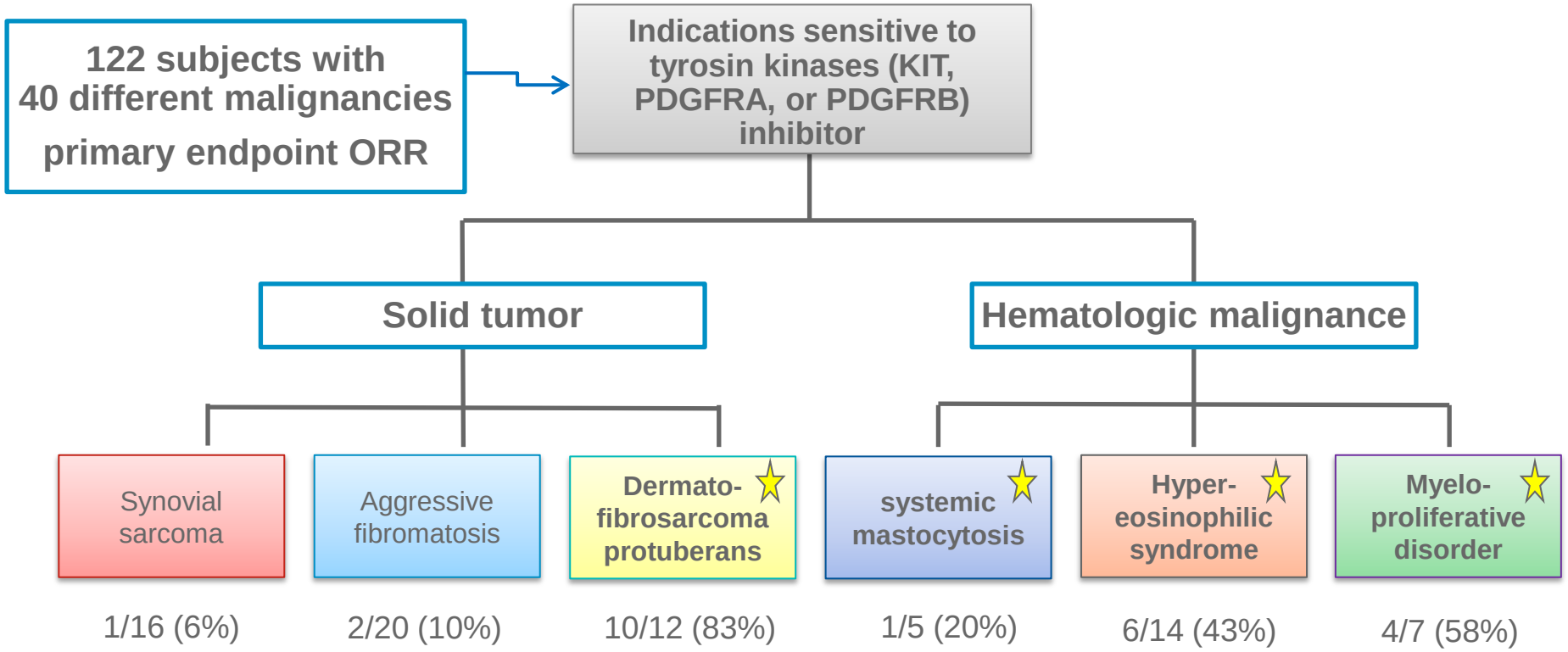


A targeted therapy focuses on a single genetic aberration and can be effective across multiple cancer types

- A large number of cancer types can be involved in the aberration
- Low frequency of the aberration
- Rarity of some of the cancers

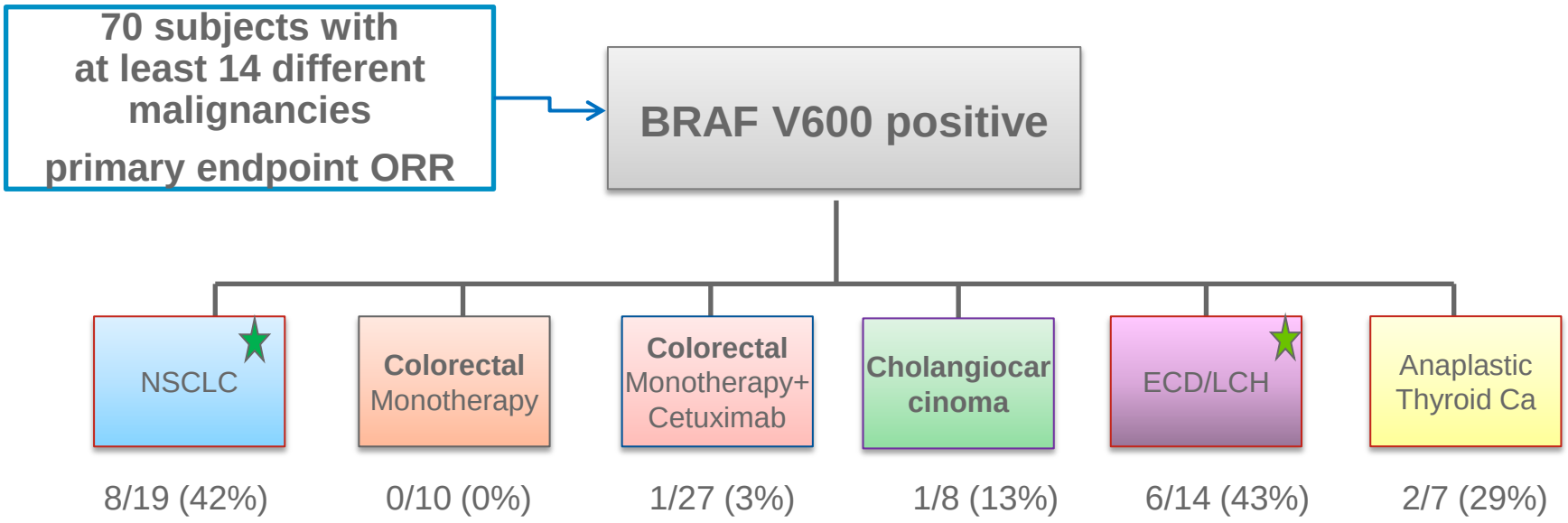
Basket trials provide an efficient tool to develop targeted cancer therapy !

First Basket Trial: Imatinib (2008)



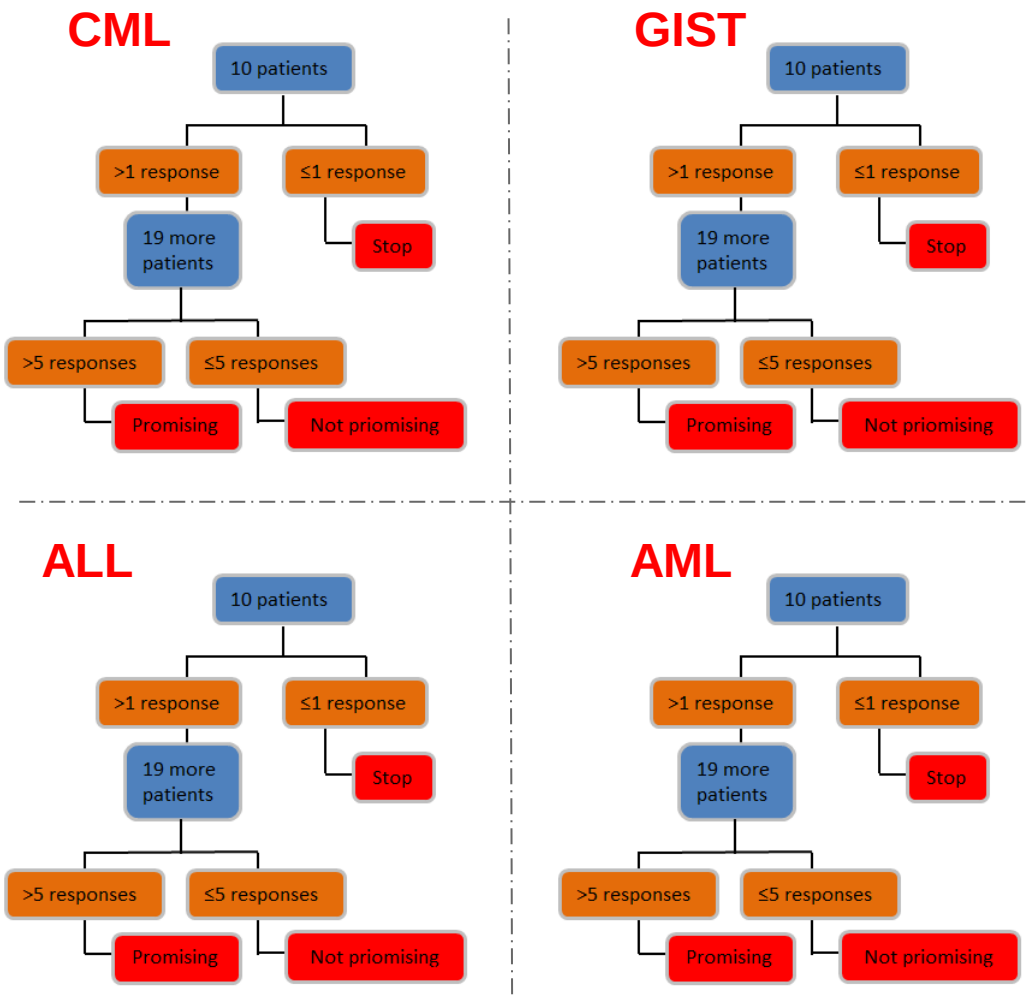
- Initially enroll up to 10 patients per cancer type
- Number of patients per indication was not prospectively stipulated
- No power consideration for sample size or inferential methods
- ★ supplemental indications after pooling from case reports and other trials

Recent Basket Trial: Vemurafenib (2015)



- An adaptive Simon two stage design was used for all tumor-specific cohorts
- No adjustment was made for multiple hypothesis testing (for false positive findings)
- Allow for additional tumor specific cohorts to be analyzed
- The histologic context is an important determinant of response in *BRAF* V600–mutated cancers.
- ★ considered to get FDA approval for these indications

Challenges in Simon Two Stage Design for Basket Trials

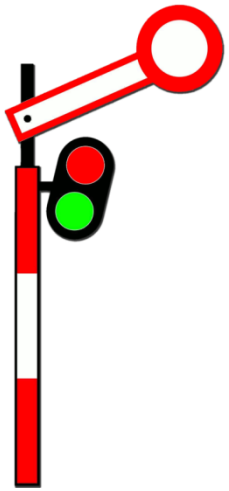


- Simon two-stage design[6]
  - Allows for stopping early due to futility
  - Distinguish a clinical meaningful response rate (30%) vs Standard of Care response rate (10%) with 5% type I error and 80% power
- Limitation of Simon two-stage parallel design in basket trials
  - Ignores the commonality among cancer type with same genetics mutation
  - Difficult with rare cancer disease

Can we do better?

Bayesian Hierarchical Modeling (BHM)

- Hierarchical modeling is a unique methodology that can be used to combine information of different indications[8, 16, 19, 20]
- Inferences for the parameters not only reflect the information about each indication, via the hierarchical modeling, but also borrow relevant information from other indications
- Sharing and borrowing information across indications allows exchangeability and improvement of power



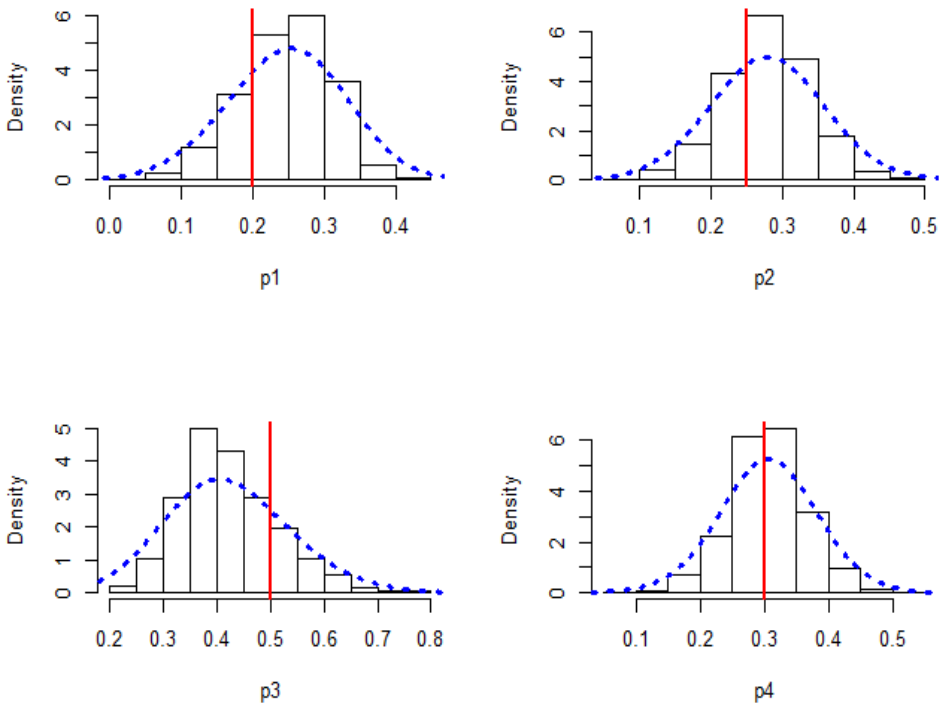
- What if there are some indication that is very dissimilar from the rest? (Nugget situation)
- The BHM will give over (or under) parameter estimates and large type one error rate (less power) due to nugget effect
- An unknown heterogeneity among indications poses a major problem

Example: Nugget Effect

Observed Response Rates:

p_1=0.2, p_2=0.25, **p_3=0.5**, p_4=0.3

Posterior Estimation BHM, Indications RR



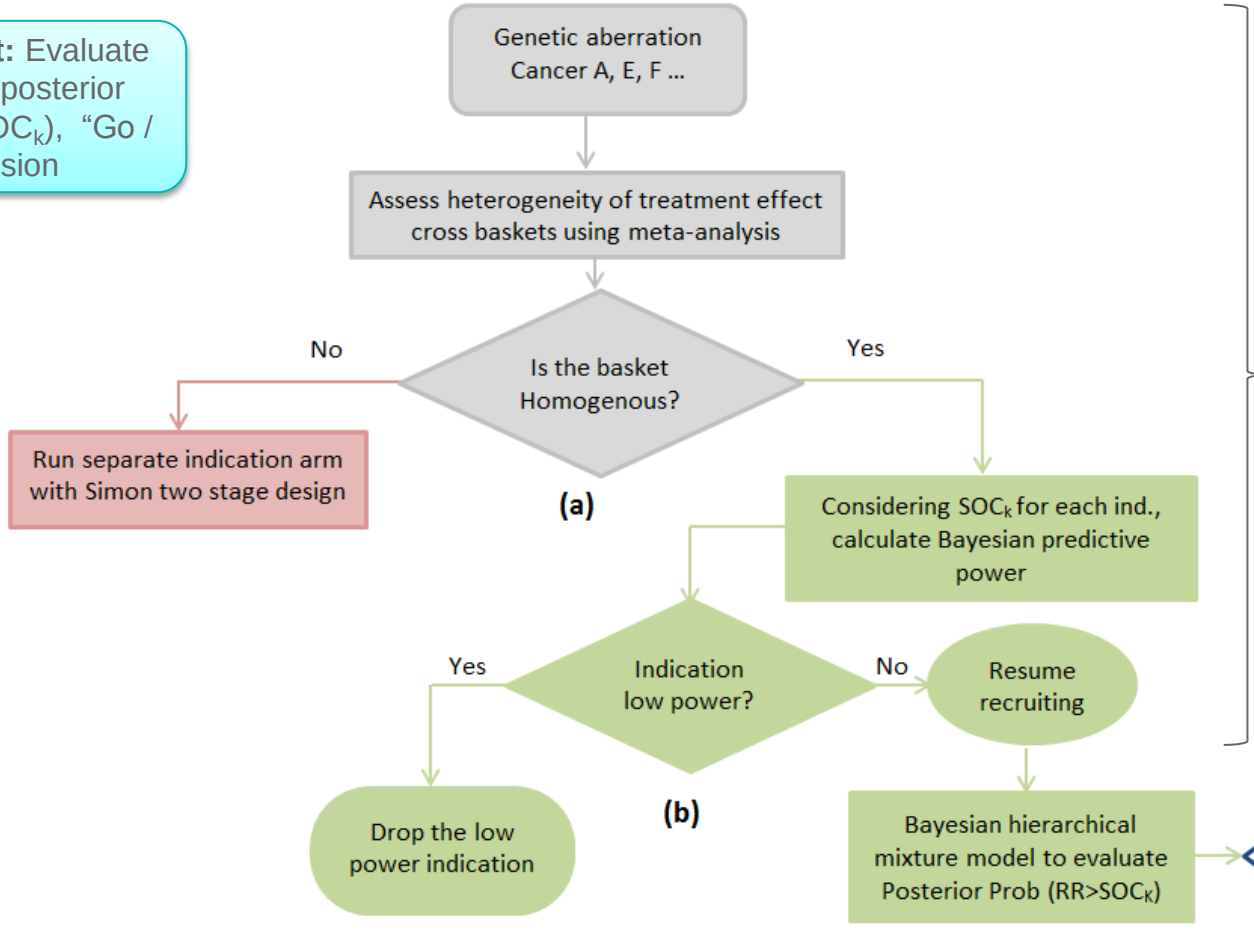
How can we avoid too optimistic/pessimistic borrowing for extreme indications (Nuggets)?

Proposed Design

Novel Design for Basket Trials using Bayesian Hierarchical Mixture Modeling(BHMM)

Proposed Novel Design for Basket Trials using BHMM

Primary Endpoint: Evaluate each indications posterior probability ($RR>SOC_k$), "Go / No-Go" decision



Proposed Design Procedures

- **Stage 1**
  - Evaluate if response rates are homogeneous across indications
    - Heterogeneous: Simon two stage parallel design
    - Homogeneous: Bayesian predictive power evaluation

Decision 1

  - Apply Bayesian predictive power assessment for early futility rule
    - Non-promising indication, drop the indication
    - Promising indication, move to stage 2

Decision 2

- **Stage 2**
  - Continue recruiting patients for promising indications
    - Determine “Go/No-go” decision using BHMM

Decision 3

Proposed Design Procedures

Stage 1
Initial Enrollment

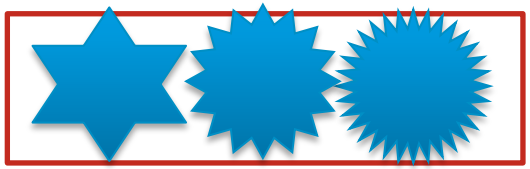
Heterogeneity Test
(P value≥0.20, Homogeneous)

RR1-RR4: True response rate

Bayesian Predictive Power
(Stop for futility, for example, probability of clinically meaningful (response ≥ SOC) is less than 50%)

S1-S4: Standard care response rate (based on historical data)

Stage 2
Resume Enrollment



Bayesian Hierarchical Mixture Model
(claim Go / No-Go, using posterior probability)

Heterogeneity Test to Mitigate Nugget Situation

- **Meta-analysis random effect model** to test response rate heterogeneity[10]
  - i. Test extreme low or high response rate indication
  - ii. Specific to binomial data and allows computation on exact binomial test
- **Under logistic-normal random effects model,**
  - Using maximum likelihood procedure, estimated between-study variance τ,

$$logit(p_i) \sim normal\ (\mu, \tau) \qquad p_i = i^{th}\ indication\ response\ rate$$

- **Test for Heterogeneity using Cochran's Q test**[12]**,**

$$H_0: p_i = p \qquad VS \quad H_\alpha: At\ least\ one\ response\ rate\ is\ different$$

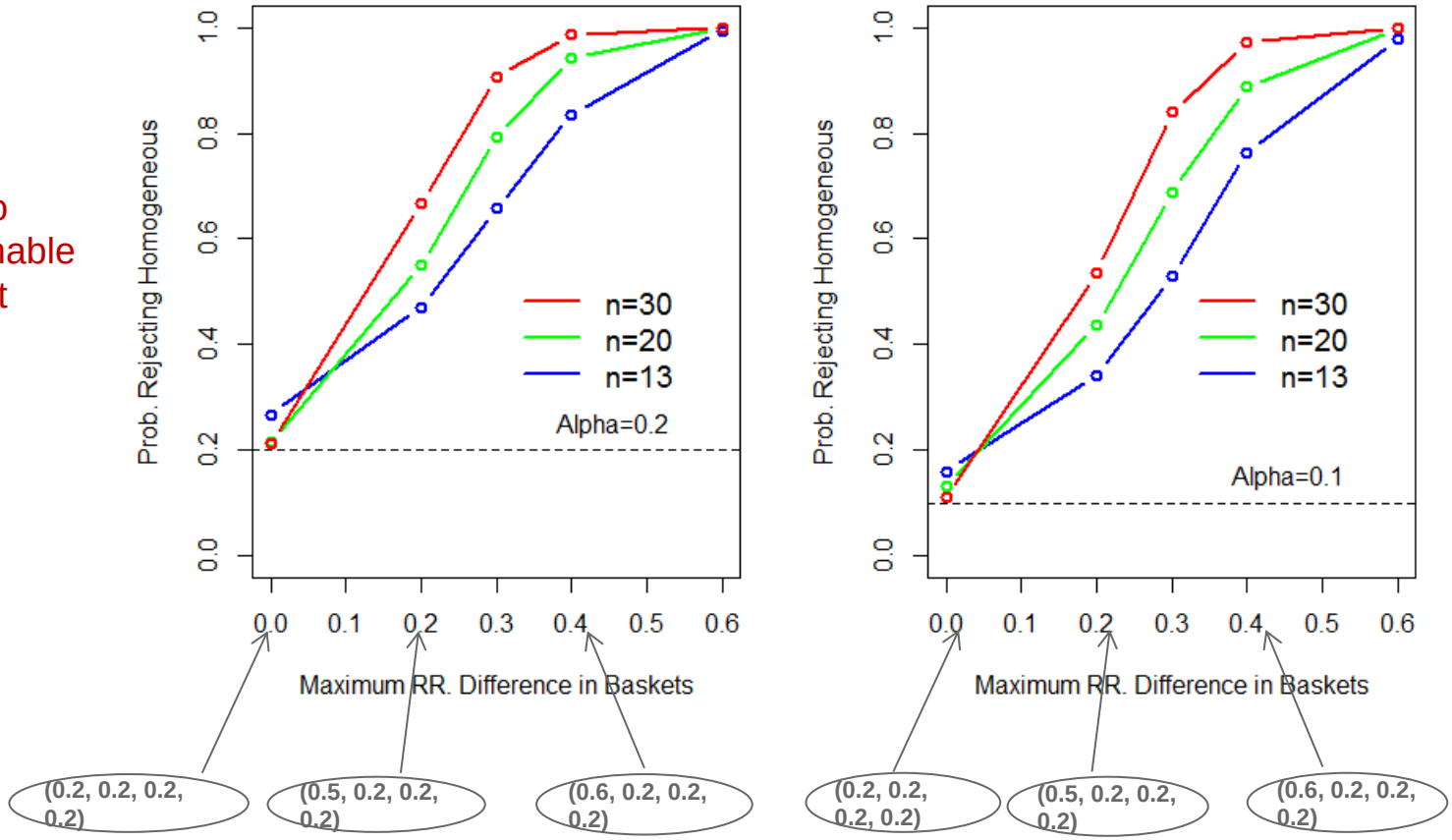
$$Q = \sum_{i=1}^{k} \frac{(\hat{p}_i - \hat{p})^2}{\tau}; \qquad where \quad \hat{p} = \frac{\sum_{i=1}^{k}(\hat{p}_i)}{k}$$

The test is conducted by comparing Q statistics to a χ^2_{k-1} distribution

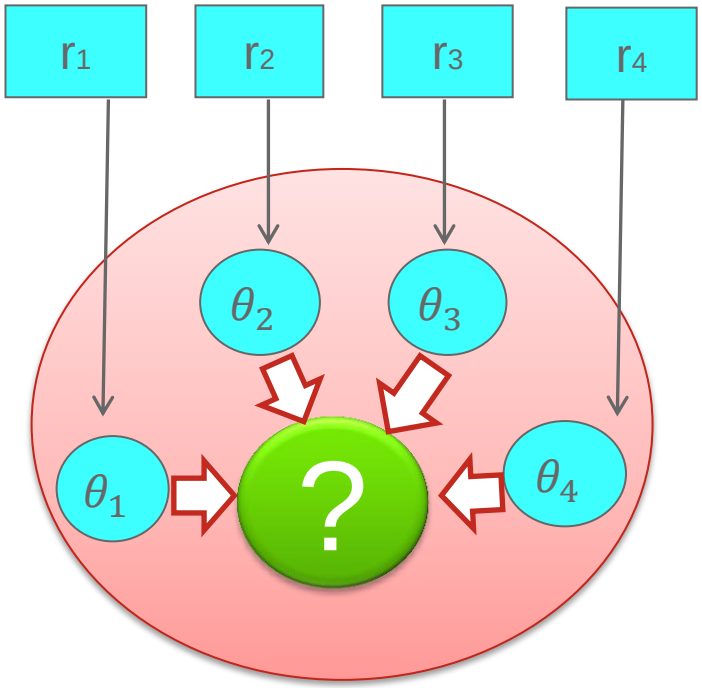
- **Decision:** If we detect heterogeneity across indications, we recommend to apply Simon two stage parallel design for each indication

Heterogeneity Test to Mitigate Nugget Situation

Tuning alpha to achieve reasonable power to detect heterogeneity



Bayesian Hierarchical Model (BHM)



Number of response:

$$r_i \sim Binomial\,(n_i, p_i), \quad i = 1, \ldots, k$$

$$logit(p_i) = \theta_i$$

First stage prior:

$$\theta_1, \cdots, \theta_k | \mu, \tau \sim N(\mu, \tau)$$

Second stage prior:

$$\mu \sim N(M, S), \qquad \tau \sim InverseGamma(\alpha, \beta)$$

μ represents indication treatment effec

τ represents variation and borrowing stre
corresponds to pooling, large τ indicate s
analyses

Can we make the prior more robust?

Bayesian Hierarchical Mixture Model (BHMM)

Mixture Prior:

- Heavy tailed mixture distribution is a robust prior[15, 17, 18]
- Gives more weight to the data when the data and the prior disagree[9]
- Share more information with observed data when they are similar. Thus achieving high precision for posterior estimation

Number of response:

$$r_i \sim Binomial\,(n_i, p_i), \quad i = 1, \ldots, k$$

$$logit(p_i) = \theta_i$$

$$\theta_i = \pi \cdot T_1 + (1 - \pi) \cdot T_2$$

π, *risk for inadequacy of prior information is constant*

T_1: *the precise information*

T_2: *the diffuse information*

First stage prior:

$$T_1 | \mu_{11,} \tau_{11}^2 \sim N(\mu_{11}, \tau_{11}^2),\ T_2 | \mu_{22,} \tau_{22} \sim N(\mu_{22}, \tau_{22}^2)$$

Second stage prior:

$$\mu_{11} \sim N(\mu_{100}, \sigma_{100}^2),\ \tau_{11}^2 \sim IG(\alpha, \beta),$$

$$\mu_{22} \sim N(\mu_{200}, \sigma_{200}^2),\ \tau_{22}^2 \sim IG(\alpha, \beta)$$

Comparison between Non-robust (BHM) versus Robust Prior Model (BHMM)

- $\theta_i \leq$ SOC$_i$ vs $\theta_i >$ SOC$_i$ for any i
- GO is defined as posterior prob$\left(\hat{\theta}_i > SOC_i | data\right) > 0.9$
- True RR: (0.30, 0.30, 0.20, 0.30), SOC: (0.20, 0.25, 0.15, 0.30)
- Simulation was based on 1000 trials

Bayesian Hierarchical Model with non-robust prior $P(\hat{\theta}_g > SOC \mid data) > 0.9$	Probability of GO for indication 1	Probability of GO or indication 2	Probability of GO for indication 3	Probability of GO for indication 4
$\boldsymbol{\mu \sim N(0.05\ 0.49)}$	0.56	0.21	0.38	0.06
$\boldsymbol{\mu \sim N(0.2, 0.29)}$	0.63	0.25	0.39	0.06
$\boldsymbol{\mu \sim N(0.8, 0.42)}$	0.66	0.31	0.41	0.08

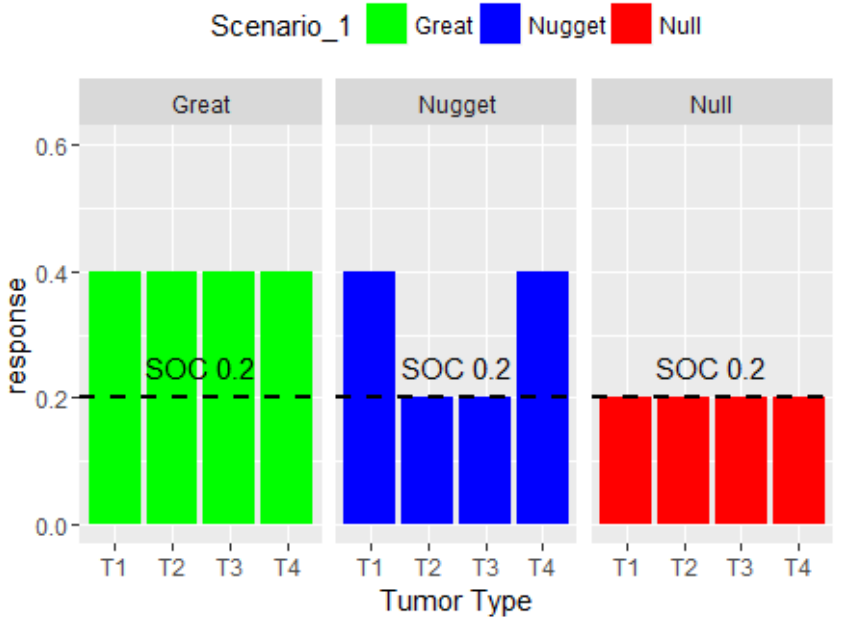
Bayesian Hierarchical Mixture Model with robust prior $P(\hat{\theta}_g > SOC \mid data) > 0.9$	Probability of GO for indication 1	Probability of GO or indication 2	Probability of GO for indication 3	Probability of GO for indication 4
$\boldsymbol{\mu_{11} \sim N(0.2, 0.29);\ \mu_{22} \sim N(0.1, 0.42)}$	0.60	0.25	0.39	0.07
$\boldsymbol{\mu_{11} \sim N(0.8, 0.42);\ \mu_{22} \sim N(0.1, 0.42)}$	0.61	0.25	0.39	0.07
$\boldsymbol{\mu_{11} \sim N(0.05, 0.49);\ \mu_{22} \sim N(0.1, 0.42)}$	0.61	0.25	0.39	0.07

Simulation Studies

- Scenario One: 4 indications in one Basket
  - Target Response Rate: (0.4, 0.4, 0.4, 0.4)
  - SOC Response Rate: (0.2, 0.2, 0.2, 0.2)
- Scenario Two: 5 indications in one Basket
  - Target Response Rate: (0.4, 0.5, 0.4, 0.4, 0.4)
  - SOC Response Rate: (0.15, 0.25, 0.2, 0.2, 0.15)
- Simulation Steps:
  - Interim Analysis
  - Parameter estimation from BHMM
  - Power and sample size evaluation

Scenario 1

True Response Rate	p1	p2	p3	p4
Great	0.4	0.4	0.4	0.4
Nugget	0.4	0.2	0.2	0.4
Null	0.2	0.2	0.2	0.2
SOC Rate	soc1	soc2	soc3	soc4
SOC Equal	0.2	0.2	0.2	0.2



Simulation study setting: **(# of simulated trials=1000, # of tumor indications=4)**

1. **Great**: Target and underlying response rate for every indication match and all of them demonstrating a promising effect in comparison with their SOC

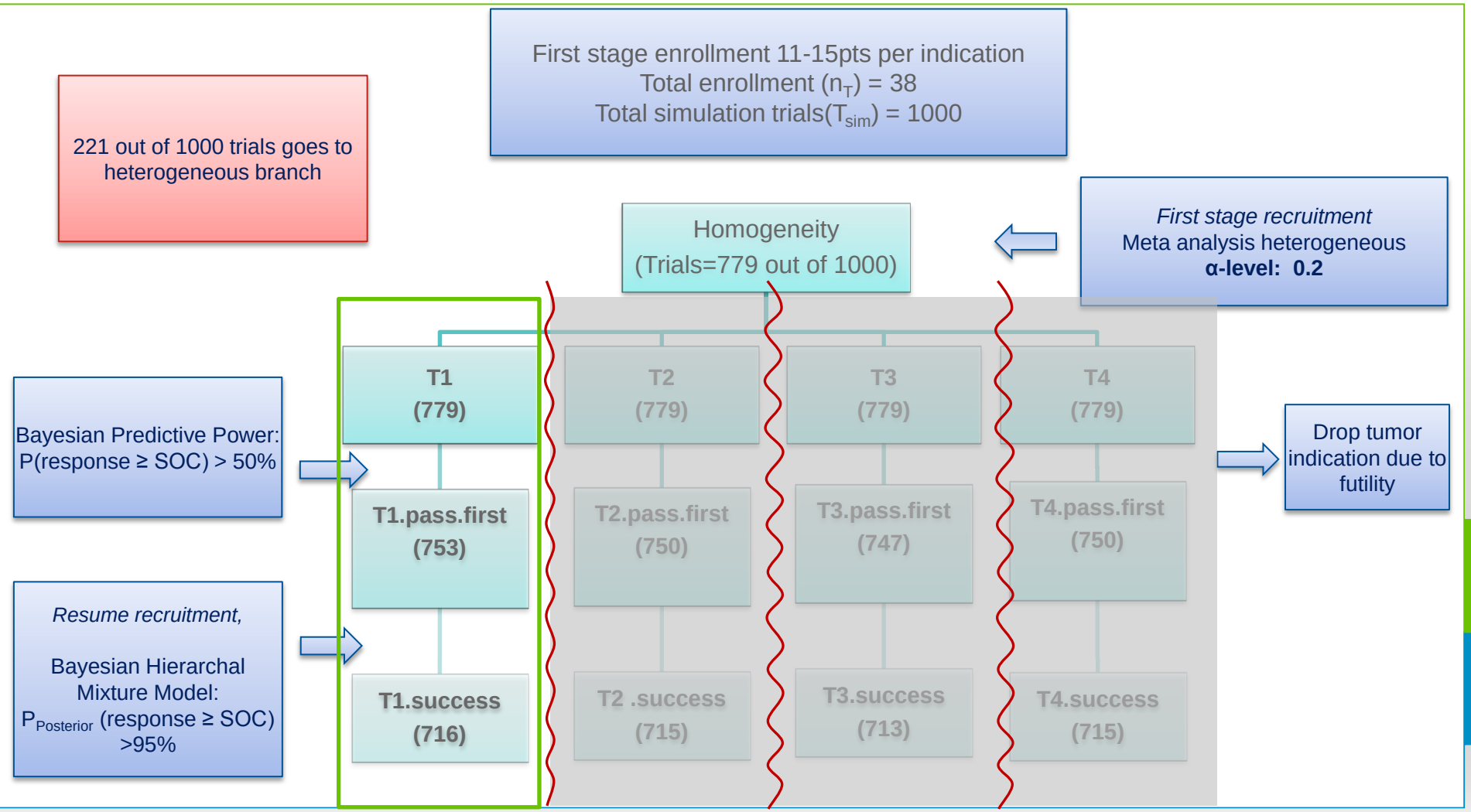
2. **Nugget**: Indication 1 and 4 similarly show promising effect, the underlying response rate for indication 2 and 3 is almost as good as SOC

3. **Null**: Every indication has an acceptable response rate but not clinically meaningful in comparison with SOC

Matching with Simon's Two Stage Design Sample Size

- Simon Two stage requires 148 patients for all four indications running in parallel to reach 80% power
- The interim analysis starts with 11 to 15 patients per indication based on Simon's two Stage Design interim criterion

Scenario 1: Study Diagram – Homogeneous Branch
Great: True=(0.4, 0.4, 0.4, 0.4) vs. SOC=(0.2, 0.2, 0.2, 0.2)



Scenario 1: Study Diagram - Homogeneous Branch - Estimations
Great: True=(0.4, 0.4, 0.4, 0.4) vs. SOC=(0.2, 0.2, 0.2, 0.2)

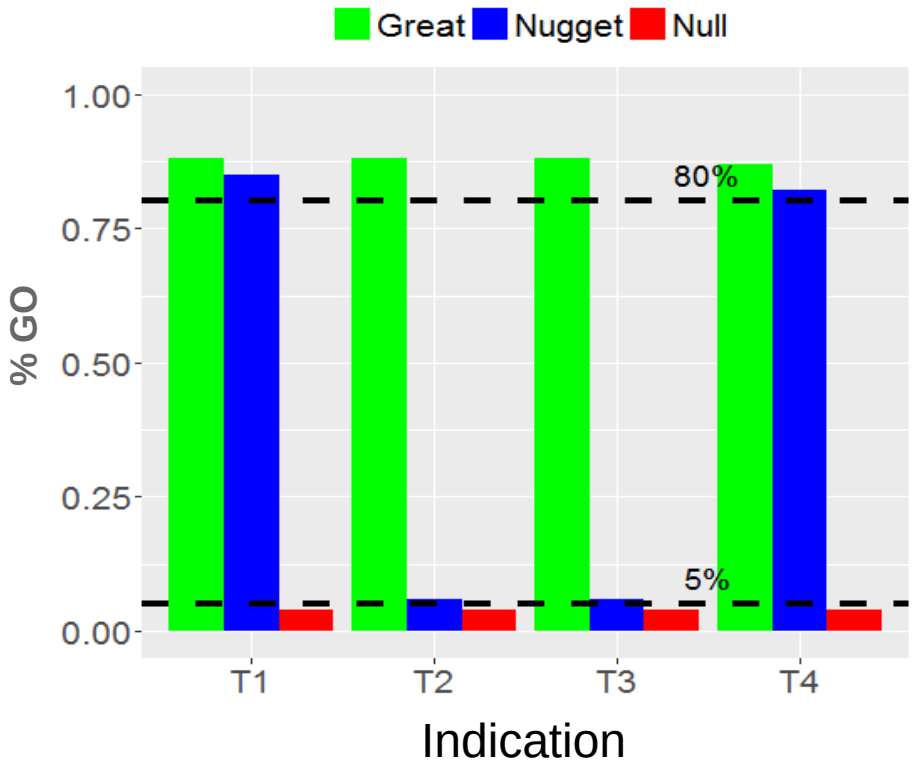
- Parameter estimation, 90% credible interval, bias, and mean squared error using Bayesian hierarchical mixture model for each indications response rate and overall response rate

Bayesian Hierarchical Mixture Model	Estimated RR1	Estimated RR2	Estimated RR3	Estimated RR4	Estimated $RR_{overall}$
Parameter Estimation True: (0.4, 0.4, 0.4, 0.4)	0.405	0.403	0.401	0.404	0.403
90% Credible Interval	(0.314, 0.502)	(0.306, 0.501)	(0.307, 0.505)	(0.309, 0.510)	(0.338, 0.469)
Bias	0.005	0.003	0.001	0.004	0.003
MSE	0.003	0.004	0.003	0.004	0.002

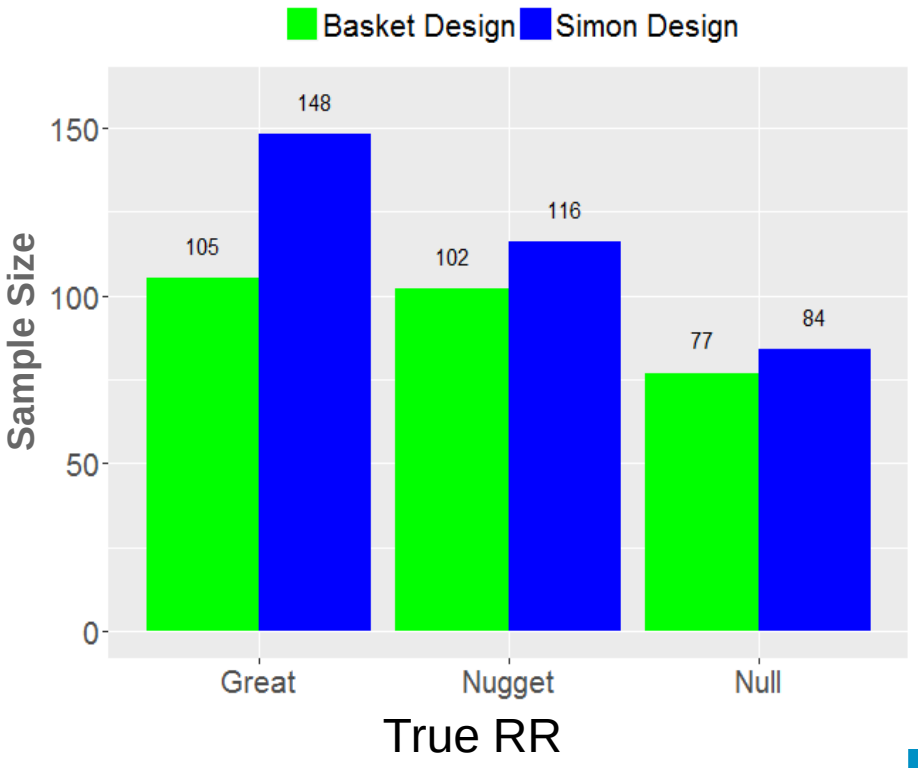
Great; True: (0.4,0.4,0.4,0.4) vs. SOC: (0.2,0.2,0.2,0.2)
Nugget; True: (0.4,0.2,0.2,0.4) vs. SOC: (0.2,0.2,0.2,0.2)
Null; True: (0.2,0.2,0.2,0.2) vs. SOC: (0.2,0.2,0.2,0.2)

Scenario 1: Simulation Final Results

% GO in Homogenous and Heterogeneous (Matching Sample Size with Simon's)



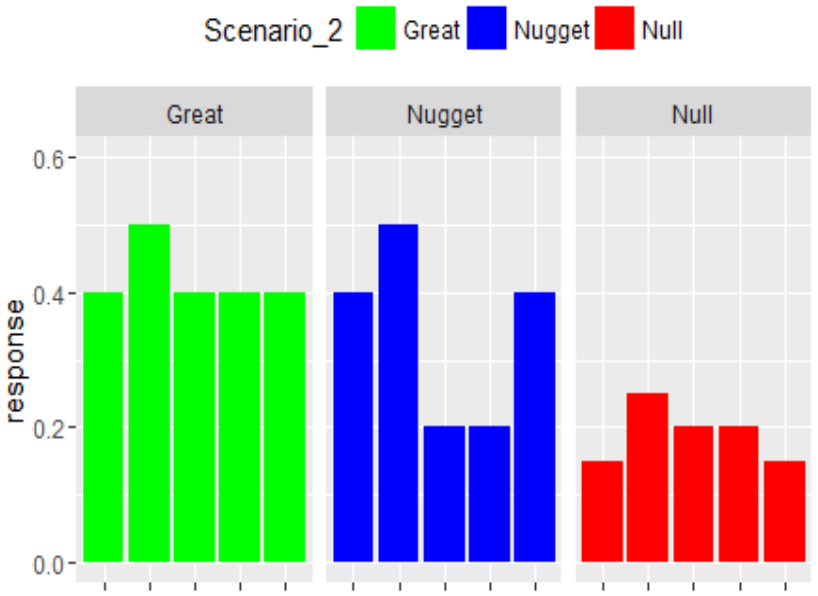
Sample Size Saving (Matching Power with Simon's)



Simon Design: 80% power and 5% α-level for each indication

Scenario 2

True Response Rate	p1	p2	p3	p4	p5
Great	0.4	0.5	0.4	0.4	0.4
Nugget	0.4	0.5	0.2	0.2	0.4
Null	0.15	0.25	0.2	0.2	0.15
SOC Rate	soc1	soc2	soc3	soc4	soc5
SOC Unequal	0.15	0.25	0.2	0.2	0.15



Simulation study setting:**(# of simulated trials=1000, # of tumor indications=5)**

1. Allow different SOC response rate across 5 indications

2. **Great**: Target and underlying response rate for every indication match and all of them demonstrating a promising effect in comparison with their SOC

3. **Nugget**: Indication 1, 2 and 5 similarly show promising effect, the underlying response rate for indication 3 and 4 is almost as good as SOC

4. **Null**: Every indication has an acceptable response rate but not clinically meaningful in comparison with SOC

Matching with Simon's Two Stage Design Sample Size

- Simon Two stage requires 138 patients for all five indications running in parallel to reach 80% power
- The interim analysis starts with 5 to 9 patients per indication based on Simon's two Stage Design interim criterion

Scenario 2: Study Diagram - Homogeneous Branch - Estimations
Great: True=(0.4, 0.5, 0.4, 0.4, 0.4) vs. SOC=(0.15, 0.25, 0.2, 0.2, 0.1)

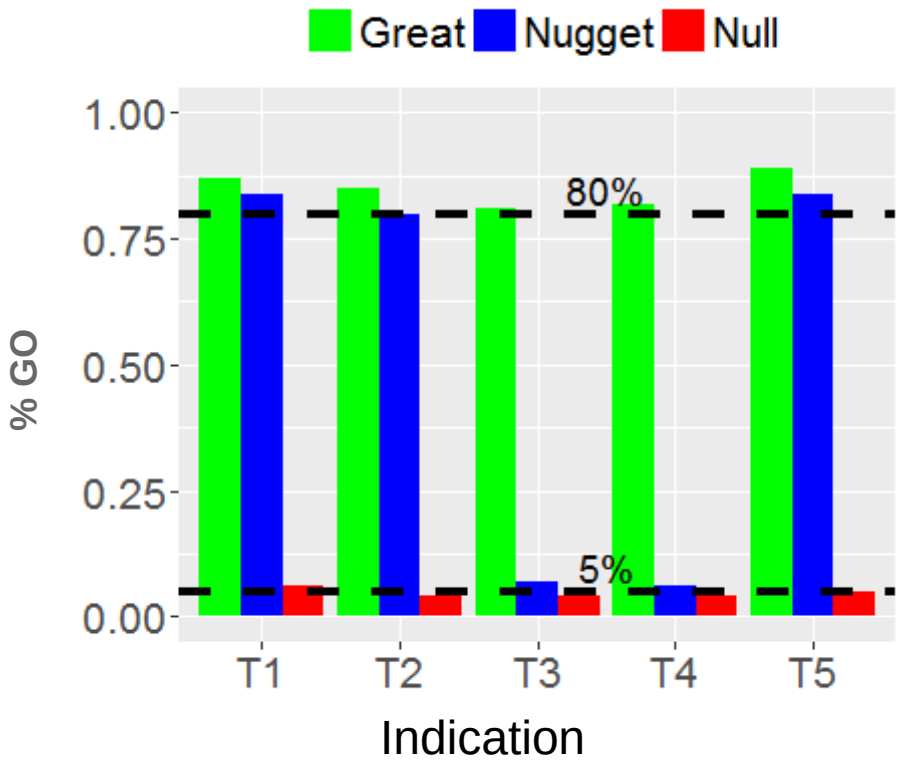
- Parameter estimation, 90% credible interval, bias, and mean squared error using Bayesian hierarchical mixture model for each indication response rate and overall response rate

Bayesian Hierarchical Mixture Model	Estimated RR1	Estimated RR2	Estimated RR3	Estimated RR4	Estimated RR5	Estimated $RR_{overall}$
Parameter Estimation True: (0.4, 0.5, 0.4, 0.4,0.4)	0.412	0.468	0.405	0.405	0.415	0.416
90% Credible Interval	(0.29, 0.55)	(0.34, 0.63)	(0.29, 0.51)	(0.29, 0.52)	(0.28, 0.55)	(0.33,0.50)
Bias	0.012	-0.032	0.005	0.005	0.015	0.001
MSE	0.004	0.006	0.003	0.003	0.005	0.002

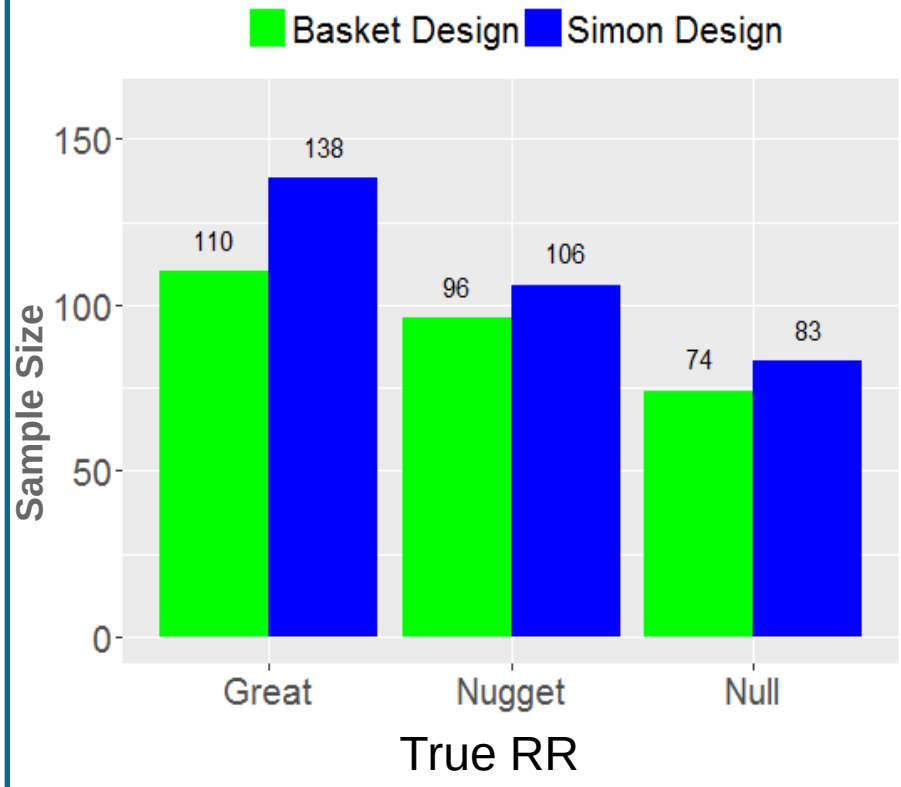
Great; True: (0.4,0.5,0.4,0.4,0.4) vs. SOC: (0.15,0.25,0.2,0.2,0.15)
Nugget; True: (0.4,0.5,0.2,0.2,0.4) vs. SOC: (0.15,0.25,0.2,0.2,0.15)
Null; True: (0.15,0.25,0.2,0.2,0.15) vs. SOC: (0.15,0.25,0.2,0.2,0.15)

Scenario 2: Simulation Results

% GO in Homogenous and Heterogeneous (Matching Sample Size with Simon's)



Sample Size Saving (Matching Power with Simon's)



Simon Design: 80% power and 5% α-level for each indication

Summary

Summary

- Histology-independent, biomarker-selected basket studies can serve as an efficient tool for developing molecularly targeted cancer therapy
- It allows for detection of early efficacy activities across multiple tumor types simultaneously
- Faster identification of efficacious drugs with fewer patients

Summary

- One challenge in interpreting the results of basket studies is drawing inferences from small numbers of patients
- This calls for innovative and efficient design:
  - The proposed design takes practical aspects of basket trial into consideration
  - It is robust to prior selection and allows dynamic borrowing of information
  - It naturally adjusts more borrowing effect when the indication are consistent and less borrowing when the indications are different
  - It saves sample size comparing to tradition two stage design and improve the efficiency of the trial

References

1. Siegel, Rebecca L., Kimberly D. Miller, and Ahmedin Jemal. "Cancer statistics, 2015." *CA: a cancer journal for clinicians* 65.1 (2015): 5-29.
2. Gorre, Mercedes E., et al. "Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification." Science 293.5531 (2001): 876-880.
3. O'Hare, Thomas, et al. "AP24534, a pan-BCR-ABL inhibitor for chronic myeloid leukemia, potently inhibits the T315I mutant and overcomes mutation-based resistance." *Cancer cell* 16.5 (2009): 401-412.
4. Westbrook, Carol A., et al. "Clinical significance of the BCR-ABL fusion gene in adult acute lymphoblastic leukemia: a Cancer and Leukemia Group B Study (8762)." *Blood* 80.12 (1992): 2983-2990.
5. Yilmaz, Z., et al. "Conventional and molecular cytogenetic findings of myelodysplastic syndrome patients." *Clinical and experimental medicine* 5.2 (2005): 55-59.
6. Simon, Richard. "Optimal two-stage designs for phase II clinical trials." *Controlled clinical trials* 10.1 (1989): 1-10.
7. Redig, Amanda J., and Pasi A. Jänne. "Basket trials and the evolution of clinical trial design in an era of genomic medicine." *Journal of Clinical Oncology* 33.9 (2015): 975-977.
8. Berry, Scott M., and Donald A. Berry. "Accounting for Multiplicities in Assessing Drug Safety: A Three-Level Hierarchical Mixture Model." Biometrics 60.2 (2004): 418-426
9. Dey, Dipak K., and Lea R. Birmiwal. "Robust Bayesian analysis using divergence measures." Statistics & Probability Letters 20.4 (1994): 287-294.
10. Nyaga, Victoria N., Marc Arbyn, and Marc Aerts. "Metaprop: a Stata command to perform meta-analysis of binomial data." Archives of Public Health 72.1 (2014): 1.
11. Schmidli, Heinz, Frank Bretz, and Amy Racine-Poon. "Bayesian predictive power for interim adaptation in seamless phase II/III trials where the endpoint is survival up to some specified timepoint." Statistics in medicine 26.27 (2007): 4925-4938.
12. Cochran, William G. "The comparison of percentages in matched samples." Biometrika 37.3/4 (1950): 256-266.
13. Jan. G. 2011. Bayesian and frequentist reasoning in plain English? answer. Available: http://meta.math.stackexchange.com/ questions/22 [2016, 11 August]
14. Kaiser, Andreas. “Mixture distributions as robust priors for Bayesian analysis.” Bayer global meeting presentation, Aug. 2014. Accessed [2016,11 August]
15. Diebolt, Jean, and Christian P. Robert. "Estimation of finite mixture distributions through Bayesian sampling." Journal of the Royal Statistical Society. Series B (Methodological) (1994): 363-375.
16. Berry, Donald A. "The brave new world of clinical cancer research: adaptive biomarker-driven trials integrating clinical practice with clinical research." Molecular oncology 9.5 (2015): 951-959
17. Schmidli H, Gsteiger S, Roychoudhury S, O’Hagan A, Spiegelhalter DJ, Neuenschwander B. Robust meta-analytic-predictive priors in clinical trials with historical control information. Biometrics 2014; 70:1023–1032.
18. Berger J, Berliner LM. Robust Bayes and empirical Bayes analysis with ε-Contaminated Priors, The Annals of Statistics, 1986 14, 461-486
19. Neuenschwander B, Wondel S., Roychoudhury S., Robust exchangeability designs for early phase clinical trials with multiple strata. Pharmaceutical Statistics, 2015 15: 123-134
20. 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.

Science For A Better Life

Thank you

Question?