abbvie

SAFETY SIGNAL DETECTION: FOR ONGOING CLINICAL TRIALS, UTILIZING A BAYESIAN FRAMEWORK

Benjamin Duncan

October 15, 2018

Safety Signal Detection: For Ongoing Clinical Trials, Utilizing a Bayesian Framework

The support of this presentation was provided by AbbVie. Abbvie participated in the review and approval of the content.

Benjamin Duncan is an employee of AbbVie Inc.

Safety Signal Detection: For Ongoing Clinical Trials, Utilizing a Bayesian Framework

Abstract:

A comprehensive and robust safety signal detection program, that includes a component which is based upon inferential statistical reasoning, is critical for clinical development programs. Safety signal detection involving ongoing (e.g. blinded) clinical trials have not always employed statistical tools and reasoning. Therefore, fundamental changes including the re-engineering of the safety signal detection process during the clinical development phase, are needed. Various statistical methods (utilizing a Bayesian framework) for assessing and evaluating blinded, clinical trials, adverse event data for possible safety signals, are presented and further evaluated. Results of simulations assessing the various methods and models will be presented. Additionally, this talk will provide more insight into the why, how and what of safety signal detection for clinical trials, and how statistical based reasoning can be implemented within a comprehensive safety signal detection process.

Agenda \ Outline

Motivation for Safety Signal Detection (SSD) and General Safety Paradigm Shift

- Why is change needed?
- How should change occur?
- What should change?

Signal detection: SSD for blinded clinical trials data

- Objectives of SSD Blinded Analyses
- Bayesian Application to SSD
- SSD Model Descriptions and Results (mock example)
- Simulation Results
- SSD Visualization
- SSD and IND Safety Reporting

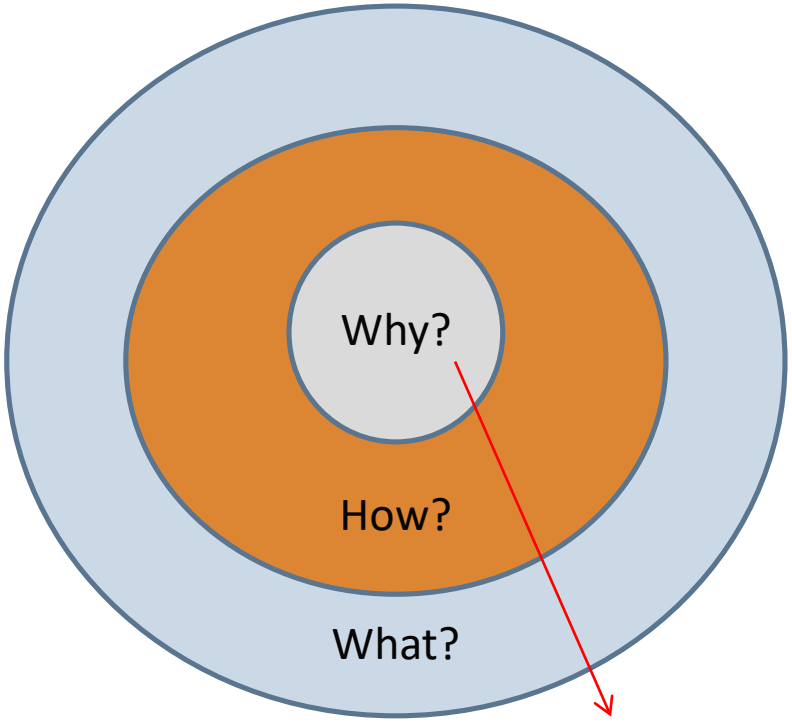
Lessons Learned / Final Thoughts / Conclusion

- Additional Research and Expansion of SSD
- Lessons Learned / Final Thoughts

abbvie

Motivation for SSD and General Safety Paradigm Shift

Motivation for SSD and General Safety Paradigm Shift



WHY is Change Needed?

➢ Pharma in the news

THE UNITED STATES DEPARTMENT of JUSTICE

en ESPAÑOL

HOME ABOUT AGENCIES BUSINESS RESOURCES

Home » Office of Public Affairs

JUSTICE NEWS

Department of Justice

Office of Public Affairs

SHARE

FOR IMMEDIATE RELEASE

Monday, July 2, 2012

[redacted] to Plead Guilty and Pay $3 Billion to Resolve Fraud Allegations and Failure to Report Safety Data

Largest Health Care Fraud Settlement in U.S. History

Global health care giant [redacted] agreed to plead guilty and to pay $3 billion to resolve its criminal and civil liability arising from the company's unlawful promotion of certain prescription drugs, its failure to report certain safety data, and its civil liability for alleged false price reporting practices, the Justice Department announced today. The resolution is the largest health care fraud settlement in U.S. history and the largest payment ever by a drug company.

WHY is Change Needed?

- **Pharma in the News**

- Vioxx

  (http://www.nbcnews.com/id/6192603/ns/health-arthritis/t/report-vioxx-linked-thousands-deaths/#.W7urnTbQaUl)

- Tamiflu

  (http://tenpennyimc.com/2013/01/12/why-you-should-avoid-tamiflu/)

- Fines for non-compliance

  (http://www.arena-international.com/pharmaco/big-pharma-being-fined-for-non-compliance-to-pv-regulations/1529.article)

WHY is Change Needed?

➢ Regulatory Environment

- Code of Federal Regulations (CFR 312.32)
- Safety Reporting Requirements for INDs and BA/BE Studies (Final Guidance, Dec 2012)
  - An “aggregate analysis of specific events observed in clinical trials that indicate those events occur more frequently in the drug treatment group than in a concurrent or historical control group”
- Safety Assessment for IND Safety Reporting (Draft Guidance for Industry, Dec 2015)
  - Sponsors should develop a Safety Assessment Committee and a Safety Surveillance Plan.
  - Sponsors should periodically review accumulating safety data, integrated across multiple completed and ongoing studies
  - Provide a quantitative framework for measuring the evidence of an association (unexpected events) or a clinically important increase (for expected events)

WHY is change needed?

➢ Regulatory Environment

Safety Monitoring and SSD requirements (regulatory)

- **Signal Definition**
  Information that arises from one or more multiple sources (including observations or **experiments**), which suggests a new, potentially **causal association**, or a new aspect of a known association between an intervention [e.g., administration of a medicine] and an event or set of related events, either adverse or beneficial, that is judged to be of sufficient likelihood to justify verificatory action. *(CIOMS, 2010, p. 14)*

- **Signal Detection Definition**

  The act of looking for and/or identifying signals using event data from any source. *(CIOMS, 2010, p. 116)*

- **The Core of Safety Signal Detection**

  Define and assess measures of disproportionate reporting (e.g. observed / expected). Identify events exceeding a specified threshold. *(Good Pharmacovigilance Practices…, 2005)*

WHY is change needed?

➢ Industry Trends and Current Literature

- Safety Monitoring in Clinical Trials *(Yao et. al., 2013)*
  - Monitoring patient safety during clinical trials is a critical component throughout the drug-development life-cycle.
  - Statistical methods, especially those based on the Bayesian framework, are important tools to help provide objectivity and rigor to the safety monitoring process.
  - Early safety signal detection not only leads to better patient protection, but also has the potential to save development costs.
- Recommendations for Safety Planning, Data Collection, Evaluation and Reporting During Drug, Biologic and Vaccine Development: *A Report of the Safety Planning, Evaluation and Reporting Team.* (Crowe et. al., 2009)
  - Proactive early planning of safety analyses in a Program Safety Analysis Plan (PSAP) and periodic aggregate safety analyses have been recommended as standard industry practices (Crowe et. al., 2009).

WHY is Change Needed?

➢ Commitment to Patient Safety

- "Patient safety is at the heart of all we do and one of our core principles. All of our medicines undergo thorough safety monitoring and evaluation processes at every stage of a medicine's lifecycle" – Teva Pharmaceuticals

- "Mallinckrodt is committed to the safety of patients, including those in the hospital settings, and the safe use of our broad portfolio of specialty pharmaceutical products," – Hugh O'Neill, Sr. VP and President, U.S. Specialty Pharmaceuticals, Mallinckrodt.

- "Beginning with the discovery of a potential new medicine, and for as long as it is available to patients, our goal is to ensure that the benefits and risks of a medication are continuously monitored and well-understood by regulators, healthcare providers and patients." – Eli Lilly

- "Part of our responsibility as a global pharmaceutical company is to help keep the patients who take our medicines safe." – Pfizer

- "Celgene is a world leader in pioneering risk minimization techniques to deliver safe use of medicinal products." – Celgene

- "Patient safety is the top priority for Biogen and AbbVie".

WHY is Change Needed?

➢ **Corporate Principles and Values**

- ❑ Accurate characterization of a compounds safety profile is essential:
  - Patient safety
  - Valuation of compound
  - Required to provide timely and accurate information on informed consent (IC) statements and investigator brochures (IB).
  - Aggregated data across all trials is required.

- ❑ Failure to report all safety findings in a timely manner leads to injury, loss of life, loss of consumer confidence for the company / industry, as well as significant financial implications for the company.

HOW Should Change Occur?

➢ **Statistical Science, Data Visualization, Coordination**

- Increased use of Data Visualization tools
  - Interactive / drill-down capabilities
  - Forest Plots, Threshold Plots, Time to Event Plots, Hazard Plots, etc.
- Use of scientific / statistical rigor (tools) for SSD blinded and aggregated unblinded analyses
  - **AEs**, Labs, Vital Signs, ECGs
- Building of global safety databases by compound for:
  - Safety signal detection
  - Aggregated safety data analysis
- Develop Program Safety Analysis Plans

WHAT Should Change?

➢ What should companies do?

- Develop Pooled, Aggregate Safety Databases
  - Earlier in the lifecycle of the clinical development program, (e.g. don't wait until time for submission)
  - Develop standard structures and reporting templates to support SSD and various other safety reporting needs

- Safety Signal Detection
  - Develop/incorporate statistical methods for blinded analysis of clinical trials data
  - Implement data visualization (static and interactive) tools
  - Optimize outputs produced for SSD

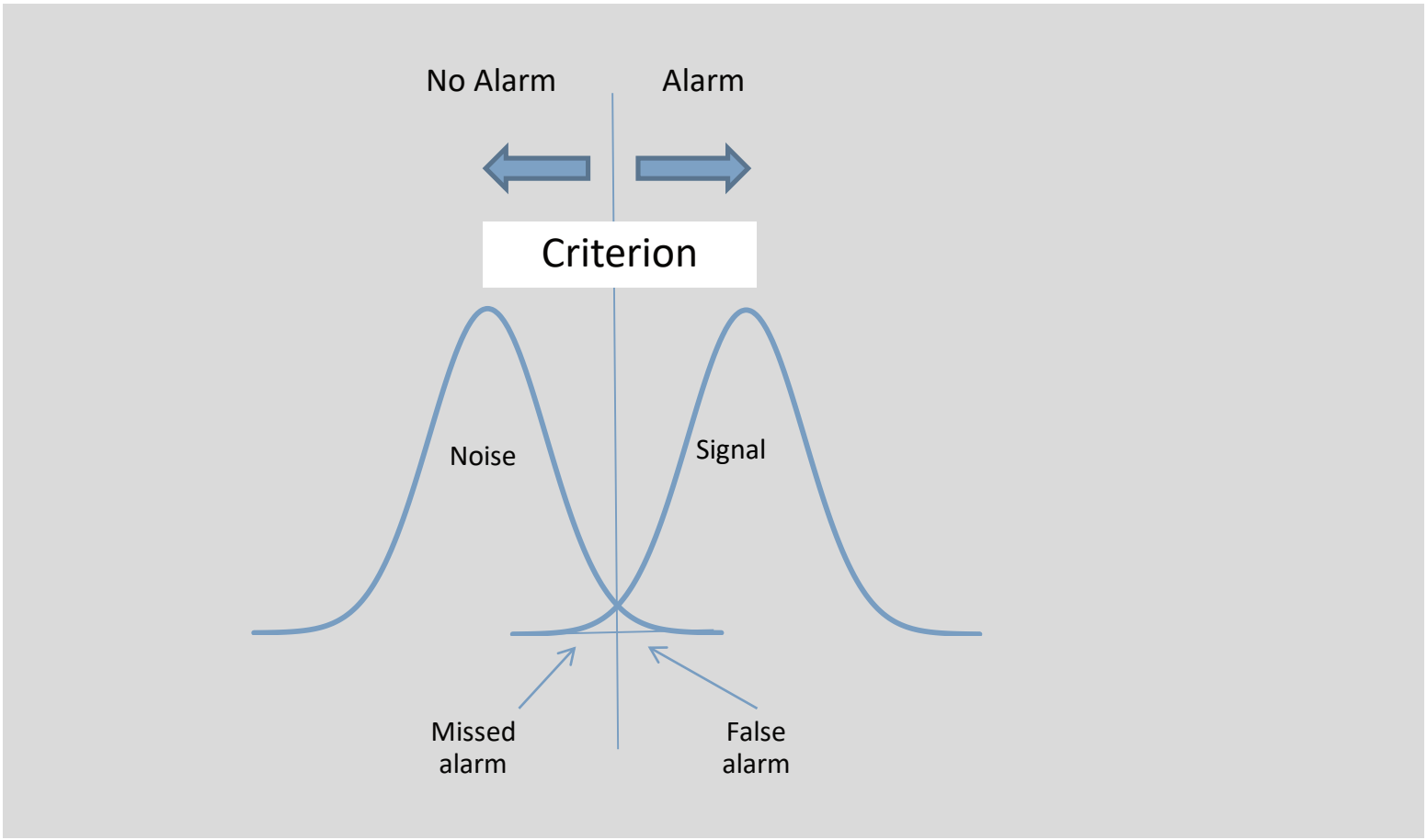
- IND Safety Reporting
  - Align the IND-SR process, around FDA guidance / regulations
  - Report SAEs that have a causal association to study drug (per sponsor's assessment, based upon medical, statistical evidence)
  - Consistent terminology for AEs/ADRs reported in IBs (e.g. Anticipated, Predicted, Expected AE of Interest, etc.)

abbvie

Signal Detection: SSD for Blinded Clinical Trials Data

Objectives of SSD Blinded Analyses

- Research for blinded analyses, Findings, and Results



Objectives of SSD Blinded Analysis

- Core Question

An Example:

- Assume that the underlying AE pbo inc. prop. for a significant event (estimated from historical data (n=500) with 24 weeks of follow-up) is 2%.

- A new blinded study has Y events after 80 subjects (3:1 randomization ratio) have completed 24 weeks.

- What is the expected value of Y, if there is no difference between actively treated subjects and current and historical placebo? How large does Y have to be to suggest that a difference (i.e. “signal”) exists?

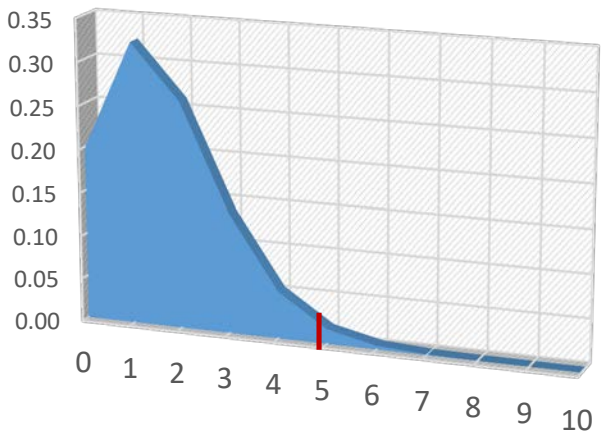
Objectives of SSD Blinded Analysis

- A Simple Frequentist Solution

Binomial Distribution

- pmf: $p(y) = \binom{n}{y} \theta^y (1-\theta)^{n-y}$, $y=0,1,2,\ldots,n$
- $E(Y_i) = \theta$, $var(Y_i) = \theta(1-\theta)$
- Y is a binomial random variable with mean and variance:
  - $\mu = E(Y) = n\theta$
  - $\sigma^2 = var(Y) = n\theta(1-\theta)$
  - Expected value: $n\theta = (80)(.02) = 1.6$

Distribution for binomial random variable Y ~ bin(n=80, θ=0.02)



n = 80, θ = 0.02

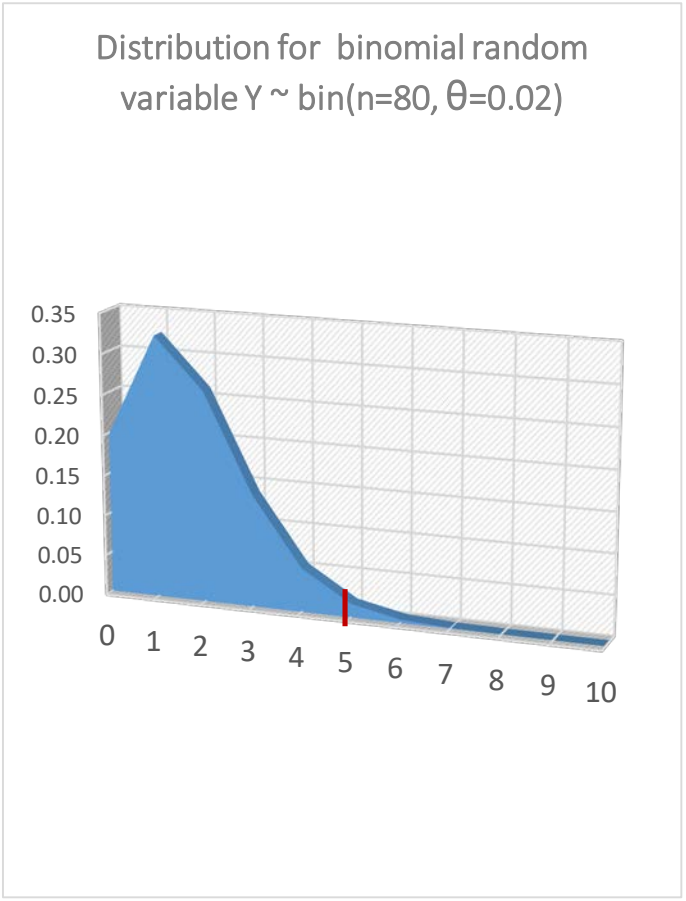
P(Y < y)	Probability
1	0.1986
2	0.5230
3	0.7844
4	0.9231
5	0.9776
6	0.9946
7	0.9989

Objectives of SSD Blinded Analysis

- A Simple Frequentist Solution: Additional Questions

Questions \ Issues

- But how did we derive our estimate for θ?
- How confident are we in our value of θ?
- What if our time-at-risk for our new blinded data is not equal to our time-at-risk from our historical data?
- What if our historical population is not representative of our new study population(s)?



n = 80, θ = 0.02

P(Y < y)	Probability
1	0.1986
2	0.5230
3	0.7844
4	0.9231
5	0.9776
6	0.9946
7	0.9989

Objectives of SSD Blinded Analysis

- How do we derive an estimate for θ?

Assume we have multiple historical studies, also assume we have constant underlying hazard rates:

(1) Compute placebo incidence rates (IR) for each study (j) for each AE of interest.

$$IR_j = \frac{r}{PT} = \frac{r}{\sum_{i=1}^{r} t_i + (n-r)T}$$

where: r = the number of patients who experience the event,

t_i = the time to event for the i^{th} patient,

n = the total number of patients

T = the average time to censor (for all patients who were censored)

→ Note that IR is the MLE of the parameter λ (the underlying hazard rate from an exponential distribution).

OBJECTIVES OF SSD BLINDED ANALYSIS

- How do we derive an estimate for θ (cont.)?

(2) Compute an overall, weighted placebo estimate for the incidence rate (e.g. the hazard rate, λ_w) for each AE of interest.

$$\lambda_w = \sum_{j=1}^{k} w_j \lambda_j = \sum_{j=1}^{k} w_j IR_j$$

where: w_j is computed from the total person-time for each study, as specified in Crowe et al (2016), [Study size adjusted method for incidence rates]

(3) Derive the expected time-at-risk distribution of patients from the blinded ongoing study (e.g. derive relative time for subjects who completed, prematurely withdrew, or were ongoing at the point of database cut-off)

OBJECTIVES OF SSD BLINDED ANALYSIS

- How do we derive an estimate for θ (cont.)?

(4) Compute expected incidence count and incidence proportion, from weighted hazard rate (step 2) and distribution of expected time-at-risk (step 3)

- Expected incidence count = E[Y] = $\sum_{i=1}^{n} y_i = \sum_{i=1}^{n}(1 - e^{-\lambda t_i})$
- Expected incidence proportion (θ) = E[Y] / n

→ θ is a time-adjusted estimate for incidence proportion

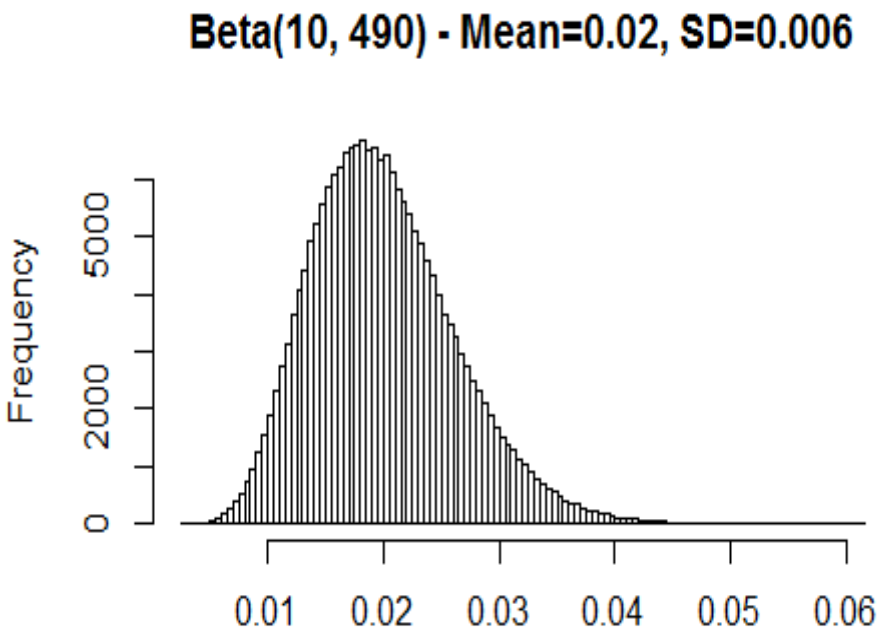
[However distribution of θ requires additional thought]

Objectives of SSD Blinded Analysis

- A Hybrid Frequentist / Bayesian Solution

Beta Distribution

- If θ is not fixed but has variation, the Beta distribution can be used to model this random variation. The Beta distribution represents a distribution of probabilities.
- pdf: $f(\theta; \alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1}(1-\theta)^{\beta-1}$, $0 \le \theta \le 1$
- Let:
  - $\mu = \frac{\alpha}{\alpha+\beta}$, $\phi = \frac{1}{\alpha+\beta}$
- Then:
  - $E(\theta) = \mu$, $var(\theta) = \mu(1-\mu)\phi/(1+\phi)$
  - Expected value (μ) = 10/(10+490) = 0.02



Percentiles: 2.5%=0.01; 25%=0.02; 50%=0.02; 75%=0.02; 97.5%=0

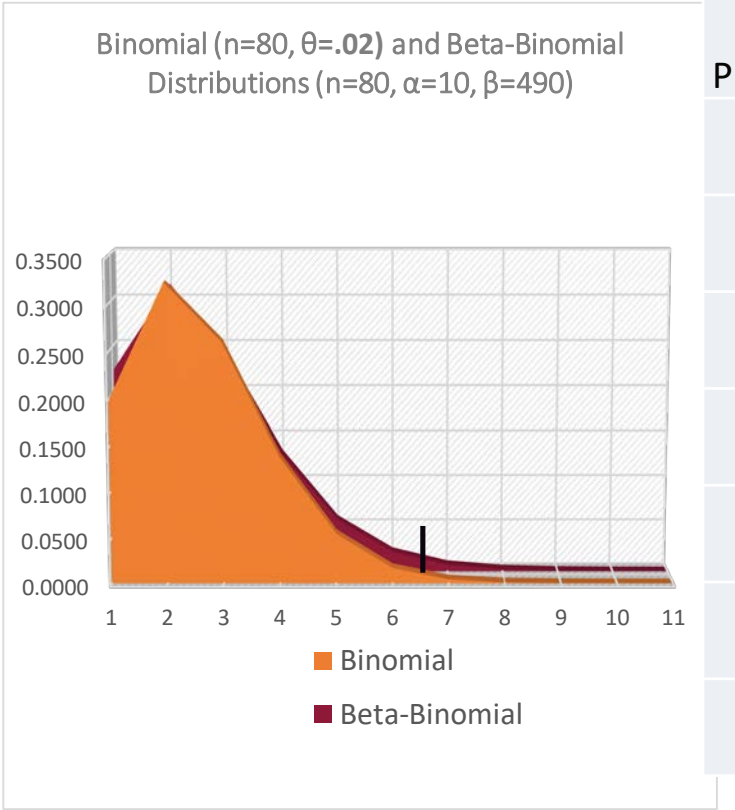
From historical data (n=500), α=10 patients with the event; β=490 patients w/o event

Objectives of SSD Blinded Analysis

- A Hybrid Frequentist / Bayesian Solution

Beta-Binomial Distribution

- Given θ, Y has a binomial distribution, bin(n, θ), and θ has a beta distribution, the resultant mixture of distributions leads to the beta-binomial distribution. Marginally, averaging with respect to the beta distribution for θ , the pmf for Y is:
- pmf: $f(y; \alpha,\beta) = \binom{n}{y} \frac{B(\alpha+y, n+\beta-y)}{B(\alpha,\beta)}$, $y = 0, 1, \ldots, n$
- Let:
  - $\mu = \frac{\alpha}{\alpha+\beta}$, $\phi = \frac{1}{\alpha+\beta}$
- Then:
  - $E(Y) = n\mu$, $var(Y) = n\mu(1-\mu)[1+(n-1)\phi/(1+\phi)]$
  - Expected value ($n\mu$) = 80(10/(10+490) = 1.6



P(Y < y)	Probability
1	0.2232
2	0.5371
3	0.7772
4	0.9093
5	0.9677
6	0.9897
7	0.9970

The Beta-binomial has slightly more spread then the binomial. However, as $\phi \to 0$, the beta-binomial converges to a binomial.

Objectives of SSD blinded analysis (Bayesian refresher)

- Bayesian Philosophy, Bayes' Theorem and Bayesian Analysis

  - Bayesian Philosophy
    - Bayes' rule provides a rational method for updating beliefs in light of new information (inductive learning ~ Bayesian inference)
    - Bayesian methods are data analysis tools derived from the principles of Bayesian inference
    - Statistical induction is the process of learning about the general characteristics of a population from a subset of members of that population.

  - Bayes' Formula/Theorem
    - P(A/B) = P(BA) / P(B) = [P(B/A)P(A)] / [P(B/A)P(A) + P(B/$\overline{A}$) P($\overline{A}$)]
    - $p(\theta|y) = \frac{p(y/\theta)p(\theta)}{\int_{\Theta} p(y/\tilde{\theta})p(\tilde{\theta})d\tilde{\theta}}$

Objectives of SSD blinded analysis (Bayesian refresher)

- Bayesian Philosophy, Bayes' Theorem and Bayesian Analysis

In Bayesian Analysis

– $P(\theta|y) \propto P(y|\theta)\,P(\theta)$

Likelihood model

Prior Distribution

Posterior Probability

- The sample space Y is the set of all possible datasets, from which a single dataset y (the observed data) will result.
- The parameter space Θ is the set of possible parameter values, from which we hope to identify the value that best represents the true population characteristics.
- The prior distribution p(θ) describes our belief that θ represents the true population characteristic (e.g. our historical data).
- The sampling model for the data p(y|θ) describes our belief that y would be the outcome of our study if we knew θ to be true.
- Once the data (y) is obtained, we update our beliefs about θ. Therefore our posterior distribution p(θ/y) describes our belief that θ is the true value, having observed dataset y.

Objectives of SSD blinded analysis (mixture models)

➢ The Form of the Finite Mixture Model

❑ Model description and specifications

- Suppose you observe realizations of a random variable Y, in which the distribution depends on an unobservable (latent) random variable S (e.g. treatment group assignment) that has a discrete distribution and can occupy one of k states (e.g. placebo, active treatment).
- Let π_j denote the probability that S takes on state j. Conditional on S = j, Y is assumed to be $f_j(y; S=j)$.
- The marginal distribution of Y is obtained by summing the joint distribution of Y and S over the states in the support of S:
  - $f(y;\alpha,\beta) = \sum_{j=1}^{k} \Pr(S = j)\, f(y; \alpha_j, \beta_j|S) = \sum_{j=1}^{k} \pi_j\, f(y; \alpha_j, \beta_j|S=j)$
  - This is a mixture of distributions and the π_j are called the mixture (or prior) probabilities.
  - This model is termed a Finite Mixture (of distributions) model, because there are k finite states of S.
  - For categorical data, the binomial and beta-binomial distributions can be specified for the data.

Objectives of SSD blinded analysis (Distribution Relationships)

- Exponential, Binomial, Poisson, Gamma

  - Exponential → Poisson
    - If the times between random events follow the exponential distribution with rate λ, then the total number of events in a time period of length t follows the Poisson distribution with parameter λt.
    - Interarrival times are independent and identically distributed exponential (λ) random variables, when λ is the rate of the Poisson process
  - Poisson → Binomial/Bernoulli
    - If we divide an interval of time into disjoint intervals of length h, where h is small [e.g. 0 –h, h-2h, 2h-3h,…], each interval corresponds to an independent Bernoulli trial, such that in each interval, there is a successful event with prob. λh.
    - Bernoulli process is a discrete time approximation to the Poisson process with rate λ, if the distribution of B(t) is approximately Poisson (λt)
  - Poisson → Gamma
    - Time until nth event occurs has a Gamma (n, λ) distribution

Objectives of SSD blinded analysis (Distribution Relationships)

- Prior Distributions for Parameters

Conjugate prior for the Exponential distribution is the gamma distribution:

$\lambda_w \sim \text{Gamma}(\lambda_w; \alpha, \beta)$

where: α is interpreted as the number of patients with the event,
β is interpreted as the total patient-time

Conjugate prior for the Binomial / Bernoulli distribution is the Beta distribution:

$\pi \sim \text{Beta}(\pi; \alpha, \beta)$

where: α is interpreted as the number of patients with the event,
β is interpreted as the number of patients without the event

Bayesian Application to SSD

- Various models / methods investigated

Table of Models

Model Description	Priors	Likelihood (data) model	SAS (Proc MCMC)	SAS (Proc FMM)	R (OpenBugs)
Pooled Prior	Overall	Binomial	Model 1	-	-
Population Mixture	Each Trt Group	Binomial	Model 2a	Model 2b	-
Individual Mixture	Each Trt Group	Binomial	Model 3a	-	Model 3b
Population Mixture	Each Trt Group	Poisson	Model 4a	-	Model 4b

Bayesian Application to SSD

- General Bayesian Framework for assessing safety signals

Framework that applies to all models

- The Bayesian framework for potential signal detection is based on evaluating the probability that a clinical parameter of interest (e.g. adverse event incidence rate or proportion) exceeds a pre-specified critical value, given the observed blinded data. Mathematically, this is formulated as an inequality around a threshold and corresponding Bayesian posterior probability and is denoted as (Wen et al., 2015):
  - **$Pr(\theta > \theta_c \mid \text{blinded observed data}) > P$ cut-off**

where: θ represents the clinical parameter of interest(e.g., "pooled blinded proportion", estimated risk difference, etc.)

θ_c represents the critical value for comparison (e.g., historical incidence proportion, or 0 if estimated risk difference is the clinical parameter of interest),

P cut-off is a probability threshold (such as 90%, 95%, or 99%) representing the desired confidence needed to identify a potential safety signal.

Bayesian Application to SSD

- Setting Priors and Thresholds

Priors

- Establishing and setting priors is critical to the final analysis. What may appear to be a non-informative prior may not be "non-informative" and could alter or bias the results.
- The neutral prior Beta(1/3, 1/3) has the unique property of centering the posterior distribution almost exactly at the sample mean (Kerman, 2011)

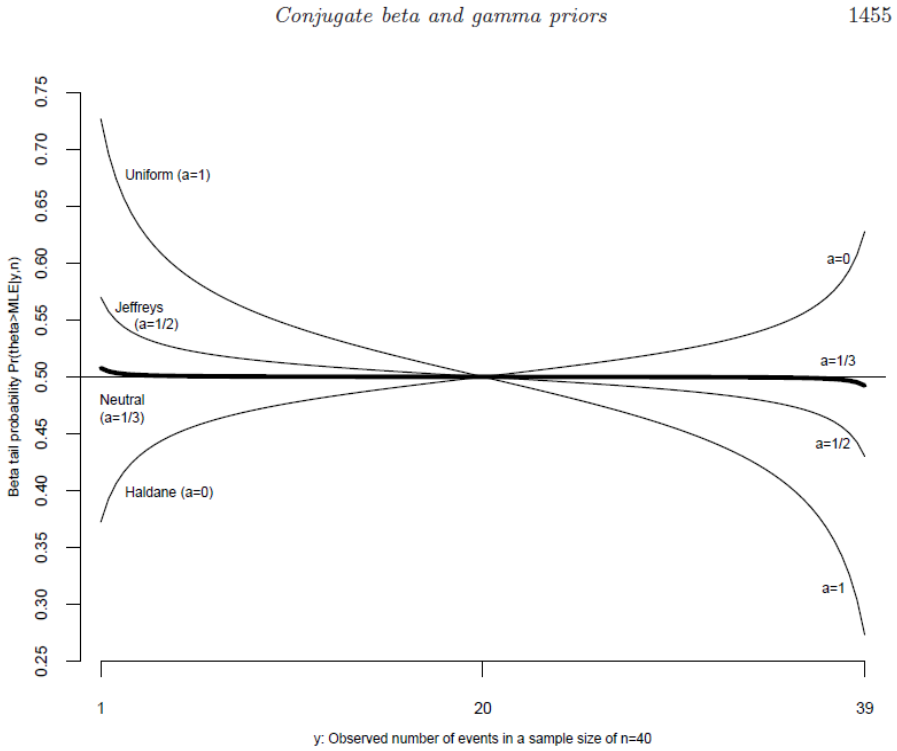


FIG 1. *Posterior tail probabilities* $\Pr(\theta > y/n|y)$ *for outcomes* $y = 1, \ldots, n-1$, *assuming a prior* $\theta \sim Beta(a, a)$ *and a sample size* $n = 40$. *The neutral prior* $Beta(1/3, 1/3)$ *has the unique property of centering the posterior distribution almost exactly at the sample mean, while other symmetric beta priors with the shape parameter* $a \leq 1$ *tend to shift the posterior mass either to the left or to the right of the point estimate, depending on the outcome.*

Source: [Kerman (2011); Electronic Journal of Statistics]

Bayesian Application to SSD

- Setting Priors and Thresholds

Priors (cont.)

- For the mixture models, the following general rule appears to have good utility:

  $\theta \sim \text{Beta}(1/3 + mp, 1/3 + m(1-p))$

  where m is the sample size and p is the incidence proportion from historical data (Kerman, 2011).

- For the overall pooled model, a neutral prior should be based upon the estimated incidence proportion from the historical placebo, so that the mean of the beta distribution equals the mean from the historical data, but down-weighted to an effective sample size of 1. [*More about this will be discussed*]

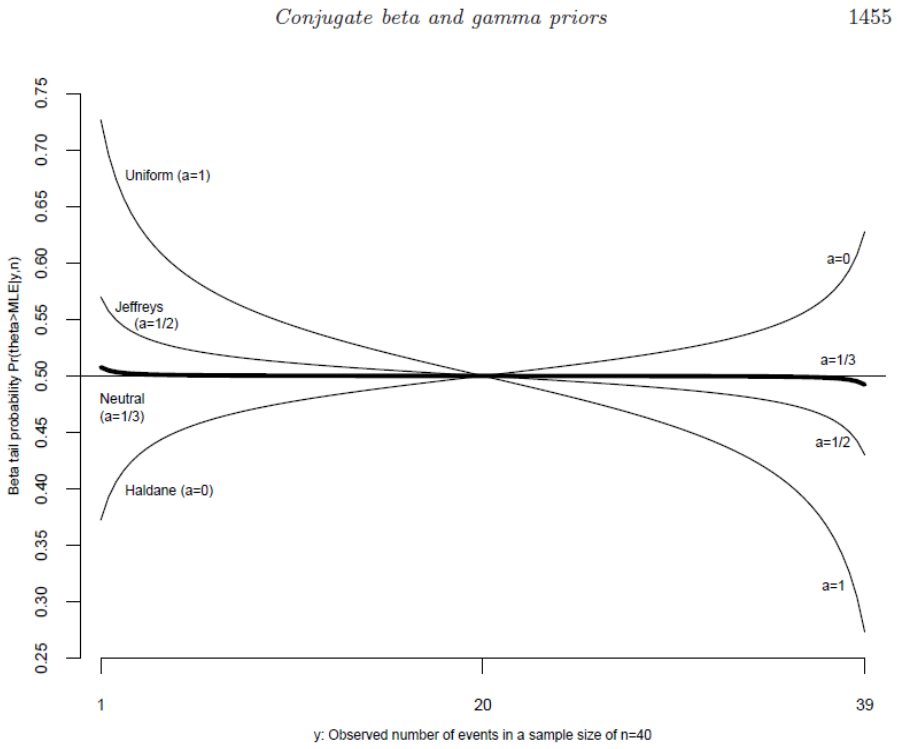


FIG 1. *Posterior tail probabilities* $\Pr(\theta > y/n|y)$ *for outcomes* $y = 1, \ldots, n-1$, *assuming a prior* $\theta \sim Beta(a, a)$ *and a sample size* $n = 40$. *The neutral prior* $Beta(1/3, 1/3)$ *has the unique property of centering the posterior distribution almost exactly at the sample mean, while other symmetric beta priors with the shape parameter* $a \leq 1$ *tend to shift the posterior mass either to the left or to the right of the point estimate, depending on the outcome.*

Source: [Kerman (2011); Electronic Journal of Statistics]

Bayesian Application to SSD

- Setting Priors and Thresholds

Thresholds

- The value for “P Cut-off” should be set to balance the sensitivity and specificity of the corresponding decision rule. Operating characteristics based upon simulations are used to assess the sensitivity (power) and specificity of the rule.

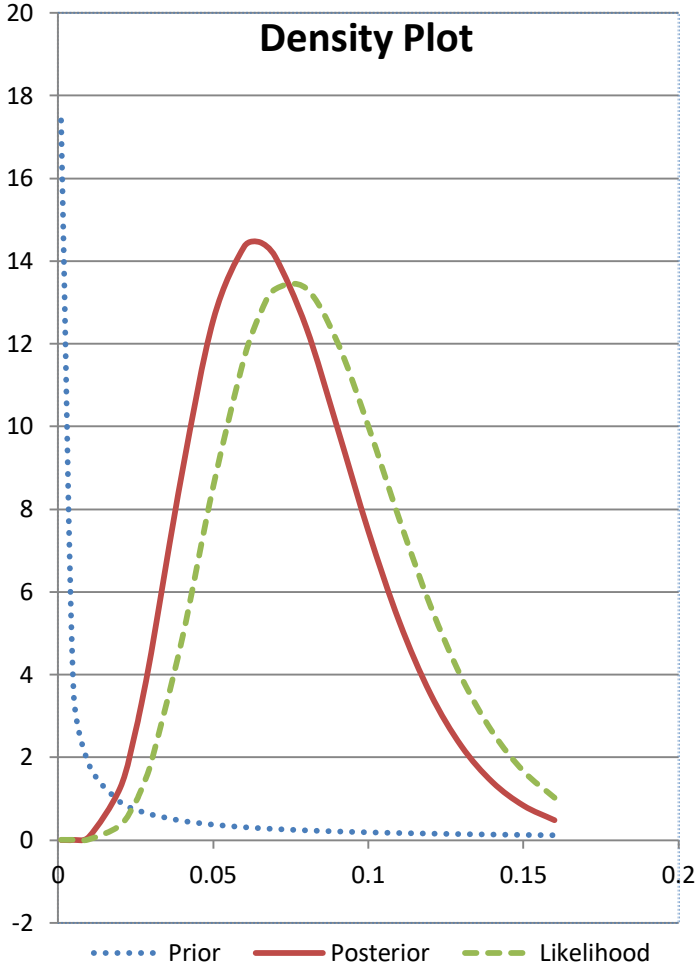
- In general, simulations suggest that less common events (e.g. incidence proportions around 5% or less) should used smaller “P Cut-off” values (e.g. 0.925 or 0.95), where events with larger incidence proportions (e.g. 10% or more) should use large “P Cut-off values (e.g. 0.975).

SSD model description and Results (mock example)

➢ Core Question Re-visited

Model 1: (Pooled Rate)

- Data: $Y = Y_{pla} + Y_{act}$ (only Y obs.)
  n = 80
- Priors: $\theta_c \sim B(\alpha,\beta)$ (for comp. only)
  $\theta \sim B(\alpha/[\alpha+\beta], 1-\alpha/[\alpha+\beta])$
- Data model: $Y \sim binom(n,\theta)$
- Results are remarkably consistent with Beta-binomial (frequentist or hybrid frequentist/Bayesian approach), which is the intent of this model.
- Bayesian framework (and Proc MCMC) offers other advantages



P(θ > θc | Y) =

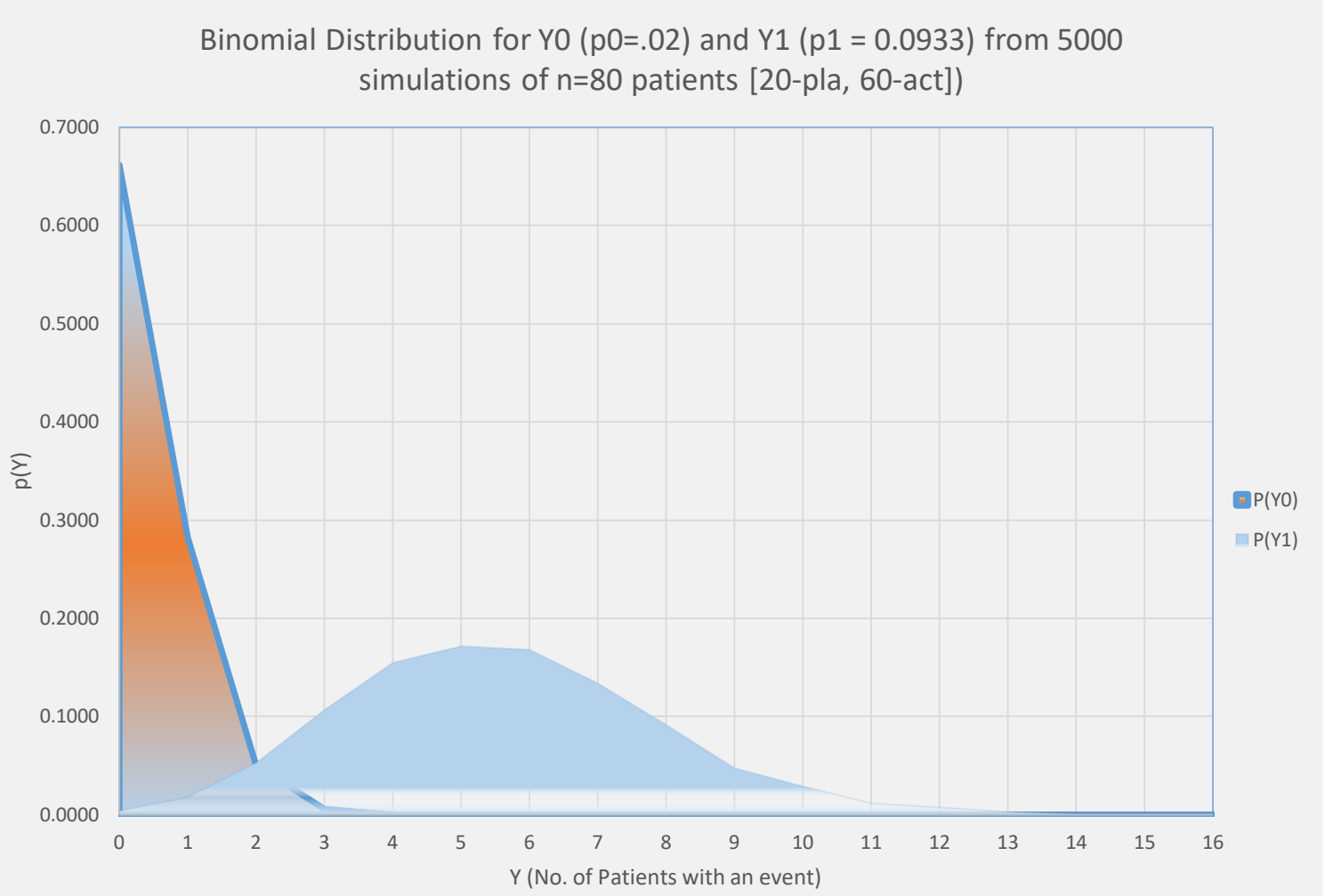
Y (data)	Posterior Probability
1	0.2258
2	0.5414
3	0.7783
4	0.9111
5	0.9687
6	0.9893
7	0.9968

Y=6 patients observed with event, n=80, Prior θ ~ B(0.02,0.98)

SSD model description and Results (mock example)

- Mixture of Binomial Distributions

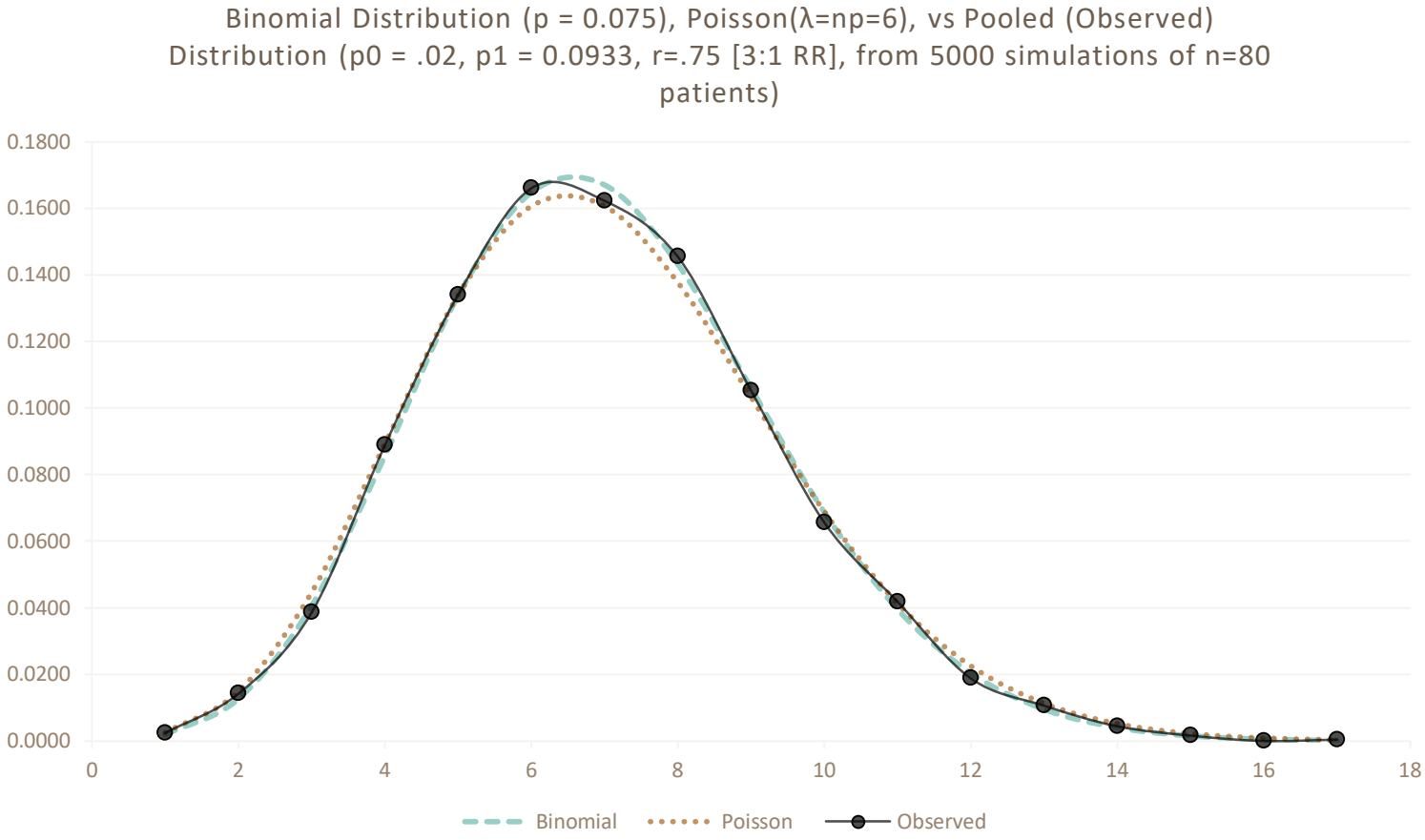
What does the resulting distribution look like when you combine different binomial distributions?



SSD model description and Results (mock example)

- Mixture of Binomial Distributions

Both the Poisson and Binomials approximate the mixture very well. [only in extreme, non-realistic settings will there be any noticeable departure]



SSD model description and Results (mock example)

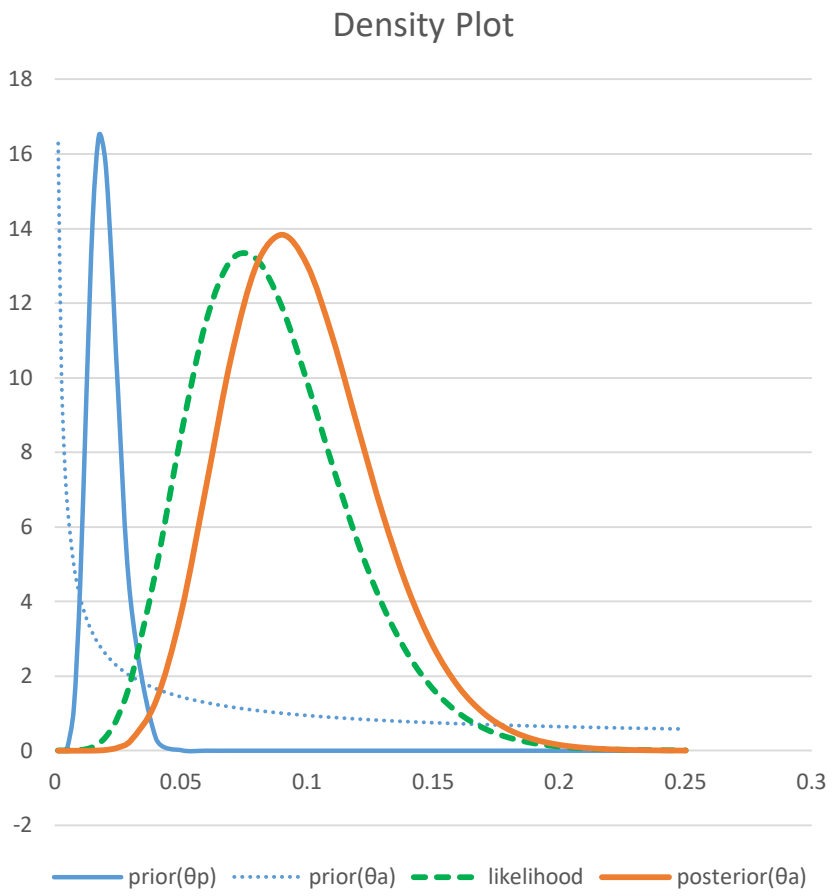
- Core Question Re-visited

Model 4a: (Pop mixture – Poisson)

- Data: $Y = Y_{pla} + Y_{act}$ (only Y obs.)
  n = 80
- Priors: $\theta_P \sim B(\alpha=10,\beta=490)$
  $\theta_A \sim B(1/3,1/3)$
  $\varphi \sim ½\ n(0,1000)$
  $P_p \sim B(200,600)$
  (rr=1:3)
- Data model: $Y \sim poisson(\lambda_y)$
  $(n-Y) \sim poisson(\lambda_{n-y})$

where: $\lambda_y = (\theta_p) \times (P_p) \times (\varphi) + (\theta_a) \times (1 - P_p) \times (\varphi)$,

$\lambda_{n-Y} = (1 - \theta_p) \times (P_p) \times (\varphi) + (1 - \theta_a) \times (1 - P_p) \times (\varphi)$



$P(\theta_A > \theta_P \mid Y) =$

Y (data)	Posterior Probability
1	0.2613
2	0.5235
3	0.7765
4	0.9156
5	0.9715
6	0.9954
7	0.9976

Y=6 patients observed with event, n=80, Prior θ_p(10,490), Prior $\theta_A \sim B(1/3 , 1/3)$

SSD model description and Results (mock example)

- Core Questions: Summary of Results from all models
  (6 events, n=80 patients, 2% historical placebo Inc. Prop. from n_h=500)

Model	θ_p	θ_a	θ	PP (Y=6)	# of events to exceed 0.975 threshold
Theoretical Expected Value (Baye's Rule)	2.0%	9.3%	7.5%	-	-
1 (Pooled Prior – Binomial)	2.0%	9.2%	7.4%	0.9893	6
2a (Pop. Mixture – Binomial [MCMC]	2.0%	9.8%	7.8%	0.9922	6
2b (Pop. Mixture – Binomial [FMM])	2.0%	9.8%	-	0.9917	6
3a (Ind. Mixture – Binomial [MCMC])	2.0%	9.7%	7.6%	0.9901	6
3b (Ind. Mixture – Binomial [R/OpenBugs])	2.0%	9.7%	-	0.9914	6
4a (Pop. Mixture – Poisson [MCMC])	2.0%	10.1%	8.1%	0.9954	6
4b (Pop. Mixture – Poisson [R/OpenBugs])	2.0%	9.8%	-	0.9918	6

SSD model description and Results (mock example)

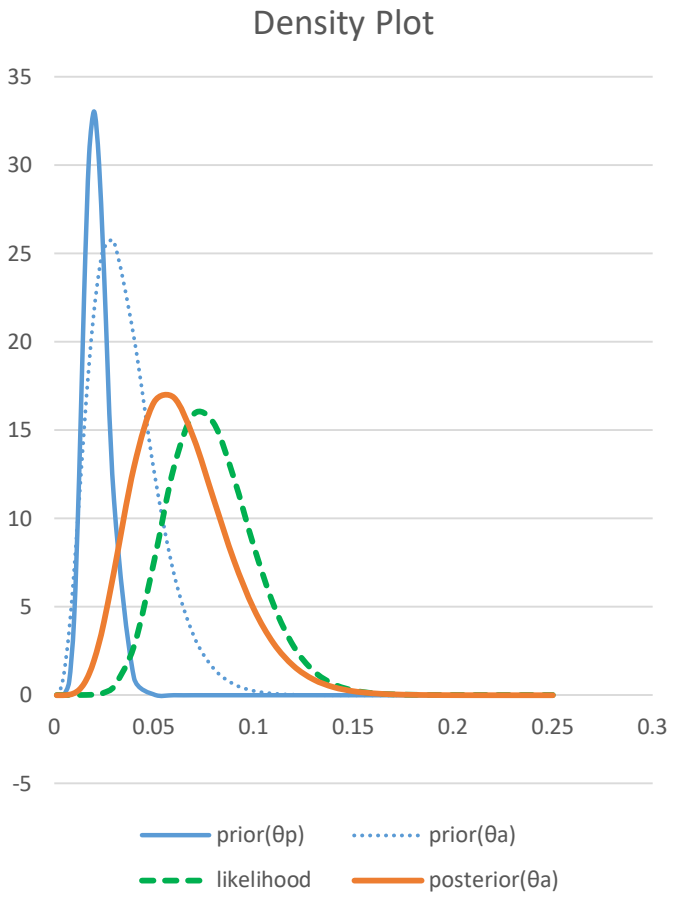
- 2nd blinded Study – 1 completed study (mock example)

Model 2a: (Binomial Pop. mixture)

- Data: $Y = Y_{pla} + Y_{act}$ (only Y obs.)
  n = 150
- Priors: $\theta_c \sim B(\alpha=10.33,\beta=490.33)$
  $\theta_{p1} \sim B(\alpha=1.33,\beta=39.33)$
  $\theta_{a1} \sim B(\alpha=4.33,\beta=116.33)$
  $P_p \sim B(200,400)$ (rr=1:2)
- Data model: $Y \sim binom(n,\theta)$

where: $\theta = (\theta_p) \times (P_p) + (\theta_{a1}) \times (1 - P_p)$

and θ_P is a weighted combination of θ_c and θ_{p1}



$P(\theta_a > \theta_p \mid Y) =$	
Y (data)	Posterior Probability
6	0.8865
7	0.9147
8	0.9404
9	0.9573
10	0.9745
11	0.9829
12	0.9900

Y=11 patients observed with event, (n=150)

SSD model description and Results (mock example)

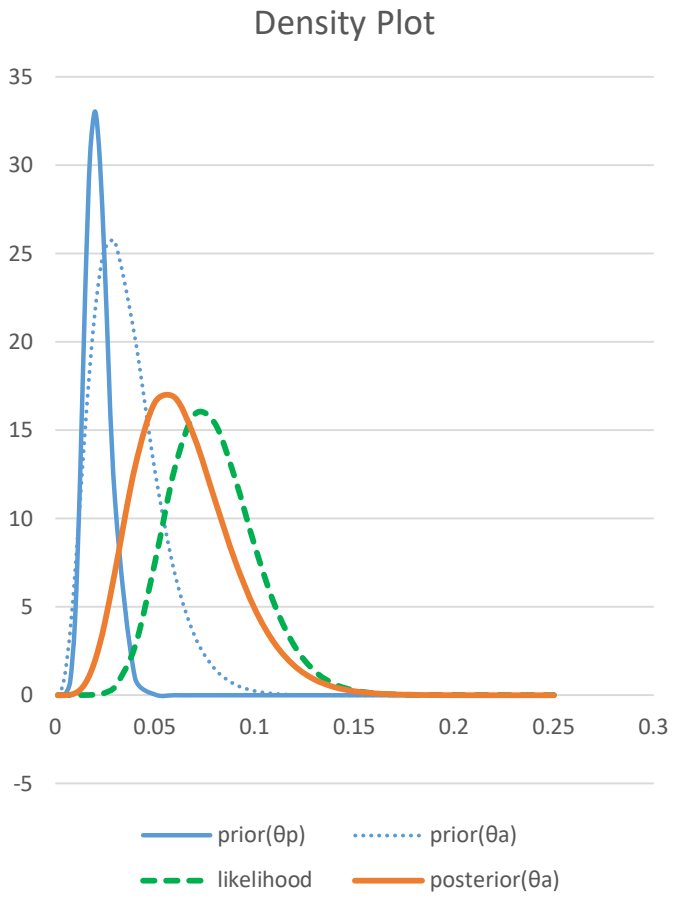
- 2nd blinded Study – 1 completed study (mock example)

Model 4a: (Pop mixture–Poisson)

- Data: $Y = Y_{pla} + Y_{act}$ (only Y obs.)
  $n = 150$
- Priors: $\theta_c \sim B(\alpha=10.33,\beta=490.33)$
  $\theta_{p1} \sim B(\alpha=1.33,\beta=39.33)$
  $\theta_{a1} \sim B(4.33,116.33)$
  $\varphi \sim$ ½ normal(0,1000)
  $P_p \sim B(200,400)$ (rr=1:2)
  θ_p = weighted comb. of θ_c and θ_{P1}
- Data model: $Y \sim poisson(\lambda_y)$
  $(n\text{-}Y) \sim poisson(\lambda_{n-y})$

where $\lambda_y = (\theta_p) \times (P_p) \times (\varphi) + (\theta_{a1}) \times (1 - P_p) \times (\varphi)$,

$\lambda_{n-Y} = (1 - \theta_p) \times (P_p) \times (\varphi) + (1 - \theta_{a1}) \times (1 - P_p) \times (\varphi)$



Y=11 patients observed with event, (n=150)

$P(\theta_a > \theta_p \mid Y) =$

Y (data)	Posterior Probability
6	0.8753
7	0.9113
8	0.9418
9	0.9601
10	0.9741
11	0.9811
12	0.9917

Simulation results (Set-up)

➢ Simulation Methods

Objective of Simulations

- ❑ Compare the various methods for efficiency and estimation accuracy
  - – Assess the sensitivity (power) and specificity (false positives) of the methods
  - – Compare estimates obtained from each model to expected values (e.g. underlying means specified in the simulated data)

Data Creation Methods Overview

- ❑ Simulated time to adverse event (using the Weibull distribution [with shape parameter=1 ~ Exponential Distribution])
- ❑ Simulated enrollment of patients using the Uniform distribution
- ❑ This allowed me to simulate separate interim SSD cuts in which patients, time-at-risk and events accrued over time.
- ❑ Number of trials = variable (up to 1,000)

Simulation results (Set-up)

- Simulation Methods

Macro variables created to control AE rates, trial duration, sample size, randomization ratios, etc.

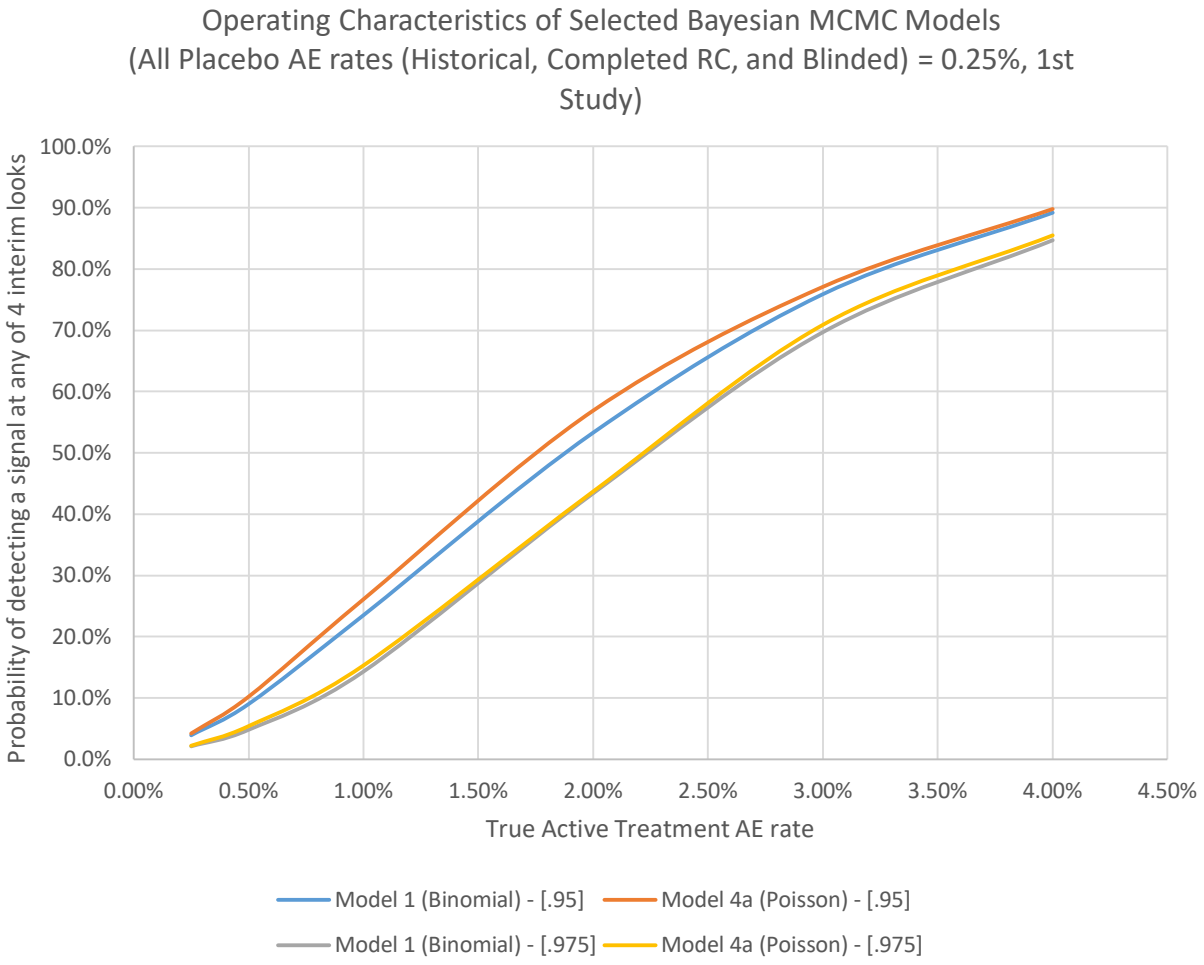
ippc	(Incidence proportion for historical placebo control at specified time)
n_hist	(historical n used for ippc, equals a+b for beta prior)
ipp	(Incidence proportion for placebo treatment, valid values: 0 - 1)
ipa	(Incidence proportion for active treatment, valid values: 0 - 1)
time	(time-at-risk [weeks])
n	(sample size per trial)
n2	(sample size per trial - 2nd trial)
ratio	(randomization ratio [active/placebo] of 1st trial
ratio2	(randomization ratio [active/placebo] of 2nd trial
trials	(number of trials to simulate)
seed	(random number generator seed)
scenario	(counter variable to keep track of scenarios and corresponding datasets)

Simulation Results (Comparing methods)

- Results

1st study – OC Curve

- Depicts the probability of detecting a signal from any of 4 interim SSD cuts.
- Both models show comparable and good specificity(low false positive rate) and comparable sensitivity (power to detect a signal, when a relevant difference exists).

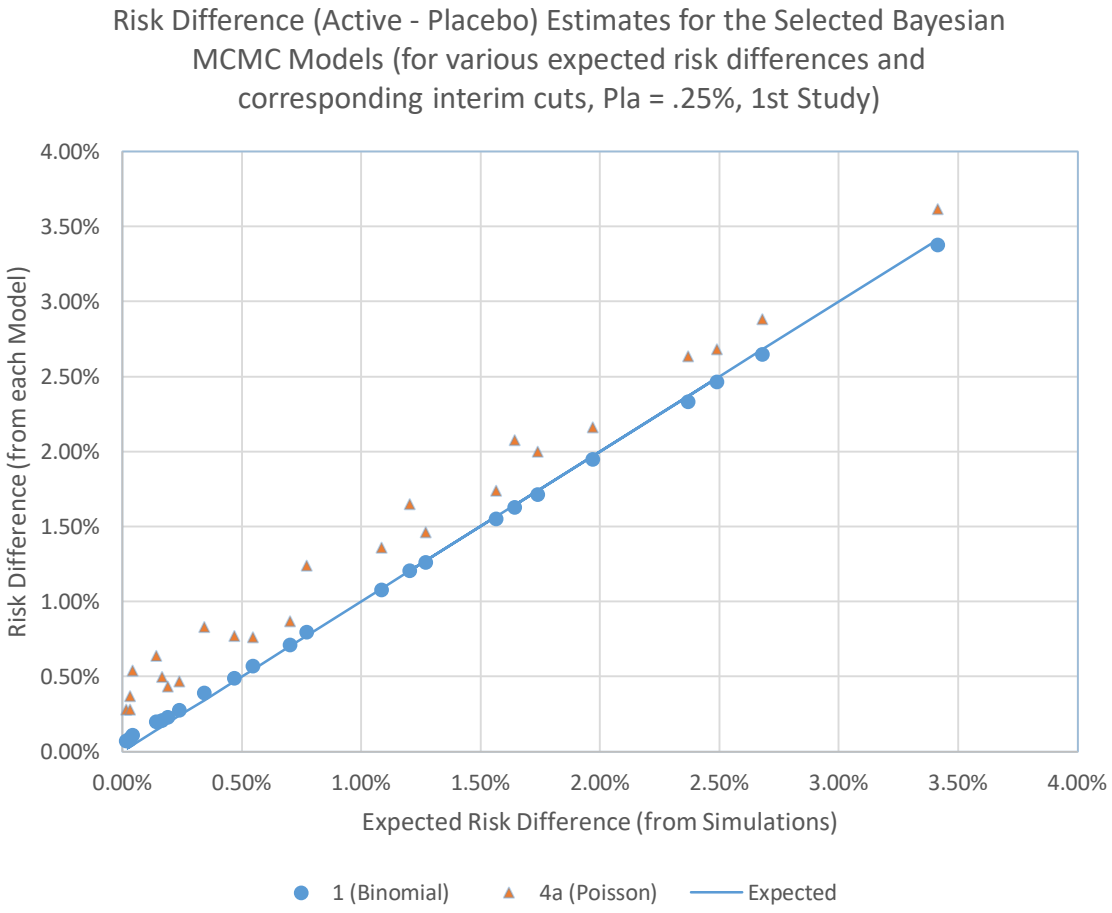


Simulation Results (Comparing methods)

➢ Results

1st study – Estimation Accuracy

- Compares risk difference from each model to the underlying true risk difference (from simulated data), for a series of 4 interims and varying levels of effect size.
- Model (1) which uses a pooled beta prior inferred from the estimated historical placebo data, and ESS=1, accurately tracks the expected result (from simulations).
- Model (4a) which uses a B(1/3, 1/3) prior for active treatment, consistently over-estimates the risk difference).

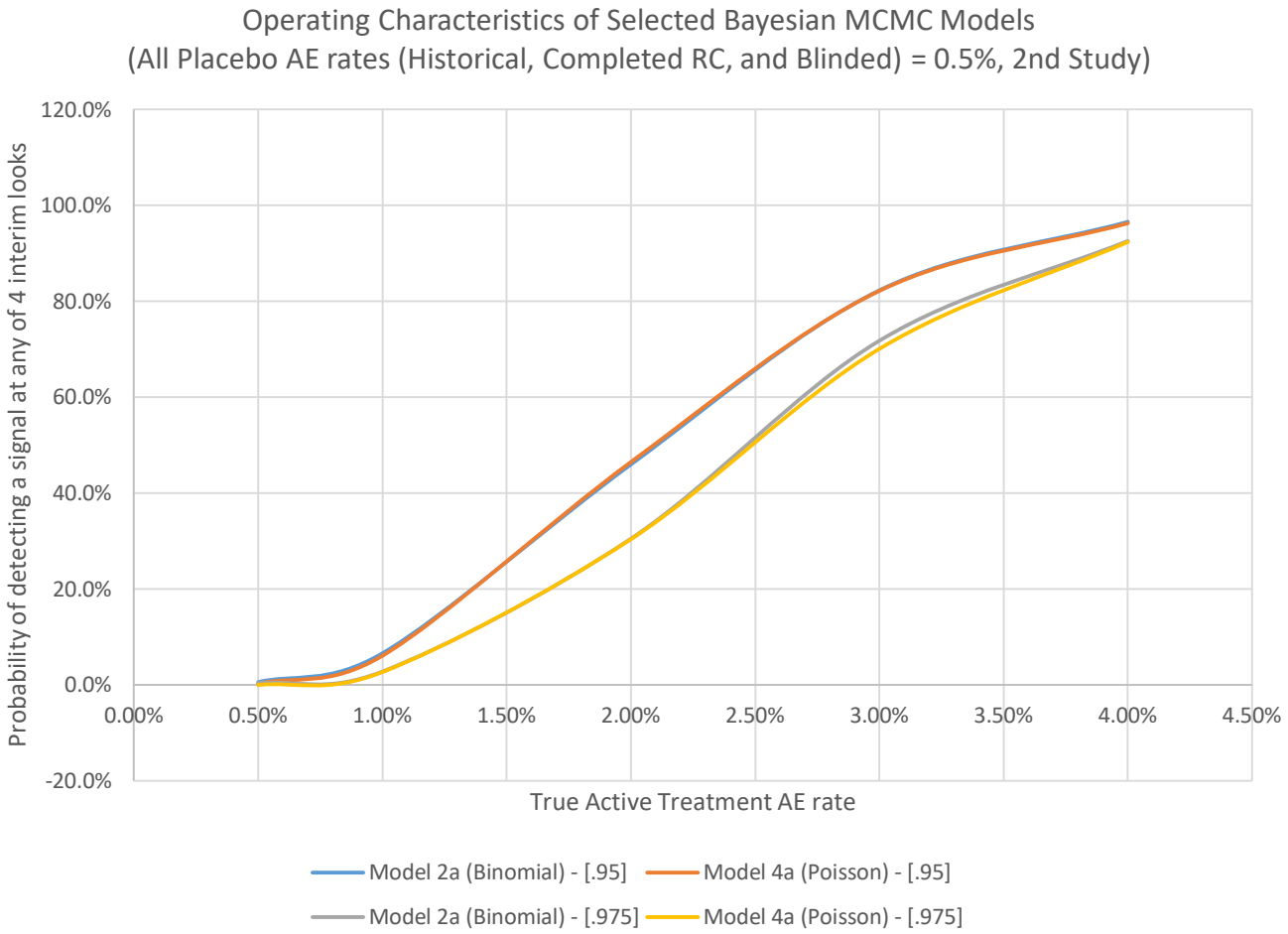


Simulation results (Comparing methods)

- Results

2nd study – OC Curve

- There is no observable difference between model 2a and 4a. They track each other perfectly.
- The specificity of both models is excellent. There is less than a 1% chance of declaring a signal, when the underlying placebo and active rates are equal.

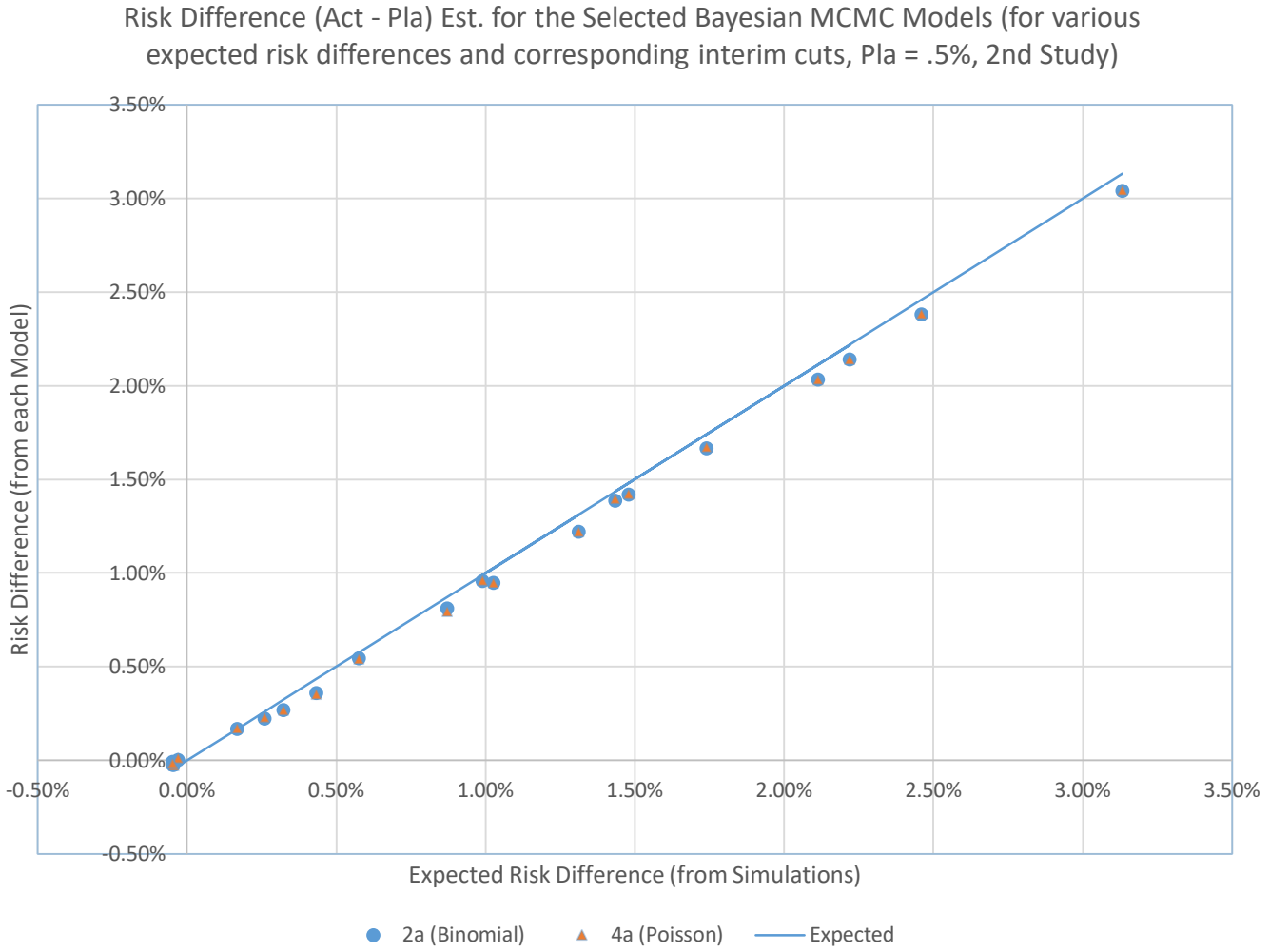


Simulation results (Comparing methods)

- Results

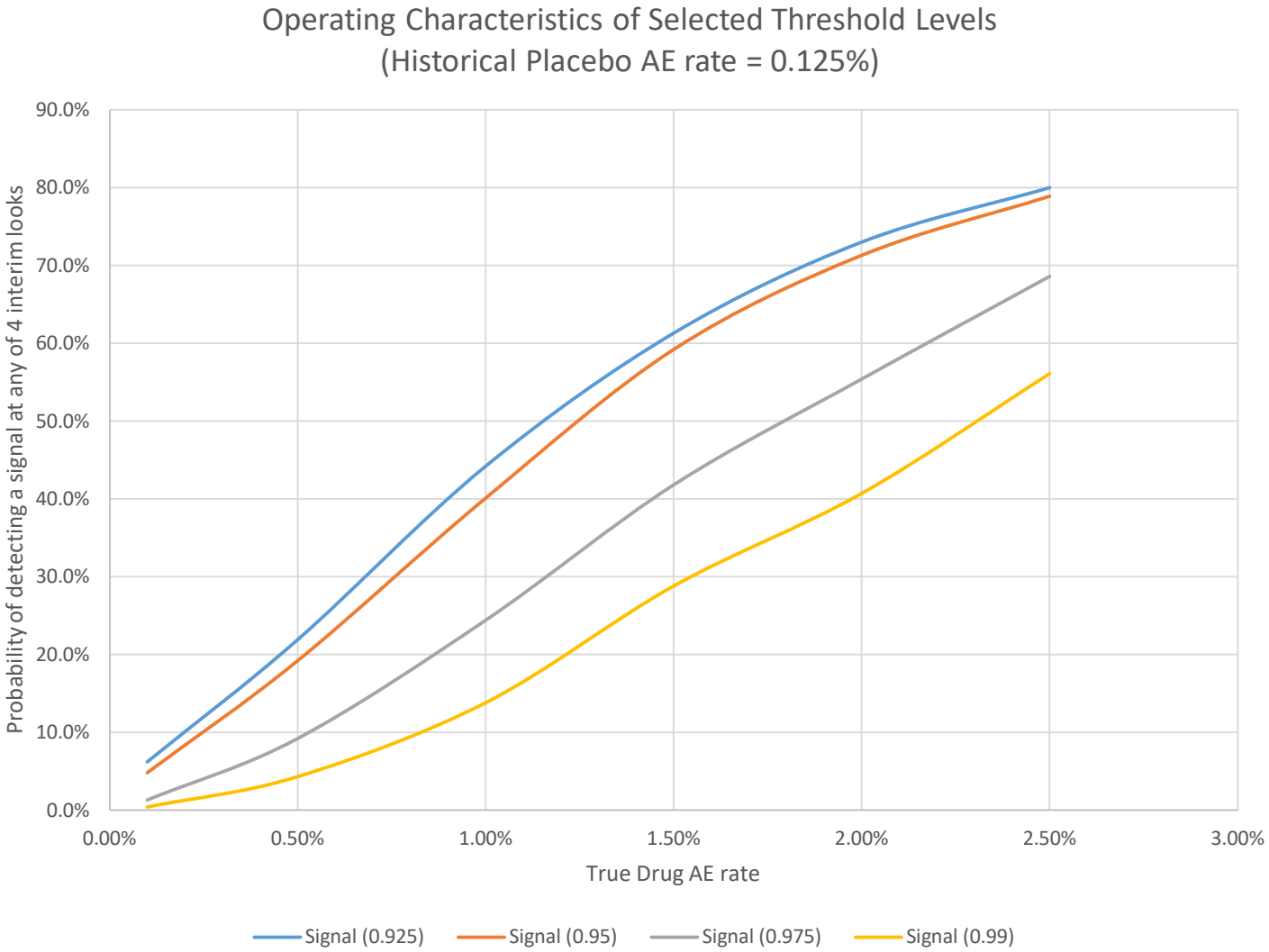
2nd study – Estimation Accuracy

- Both models very slightly under-estimate the expected results (from simulations). However, there is no observable difference between the two models.



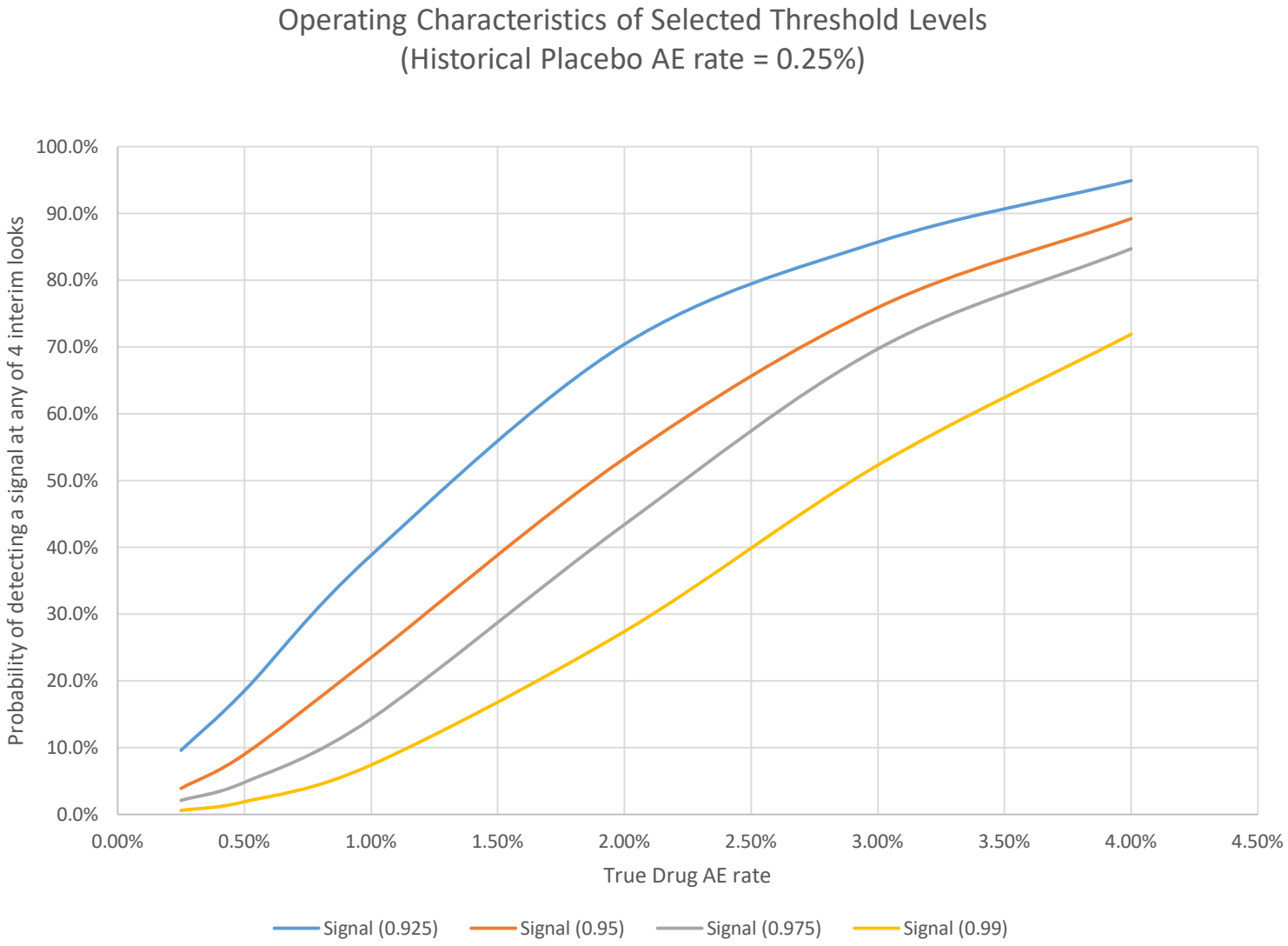
Simulation Results (comparing threshold cut-points)

False positive rates are fairly well controlled for all examined cut-points, with good sensitivity (power) for the 0.925 and 0.95 threshold levels.



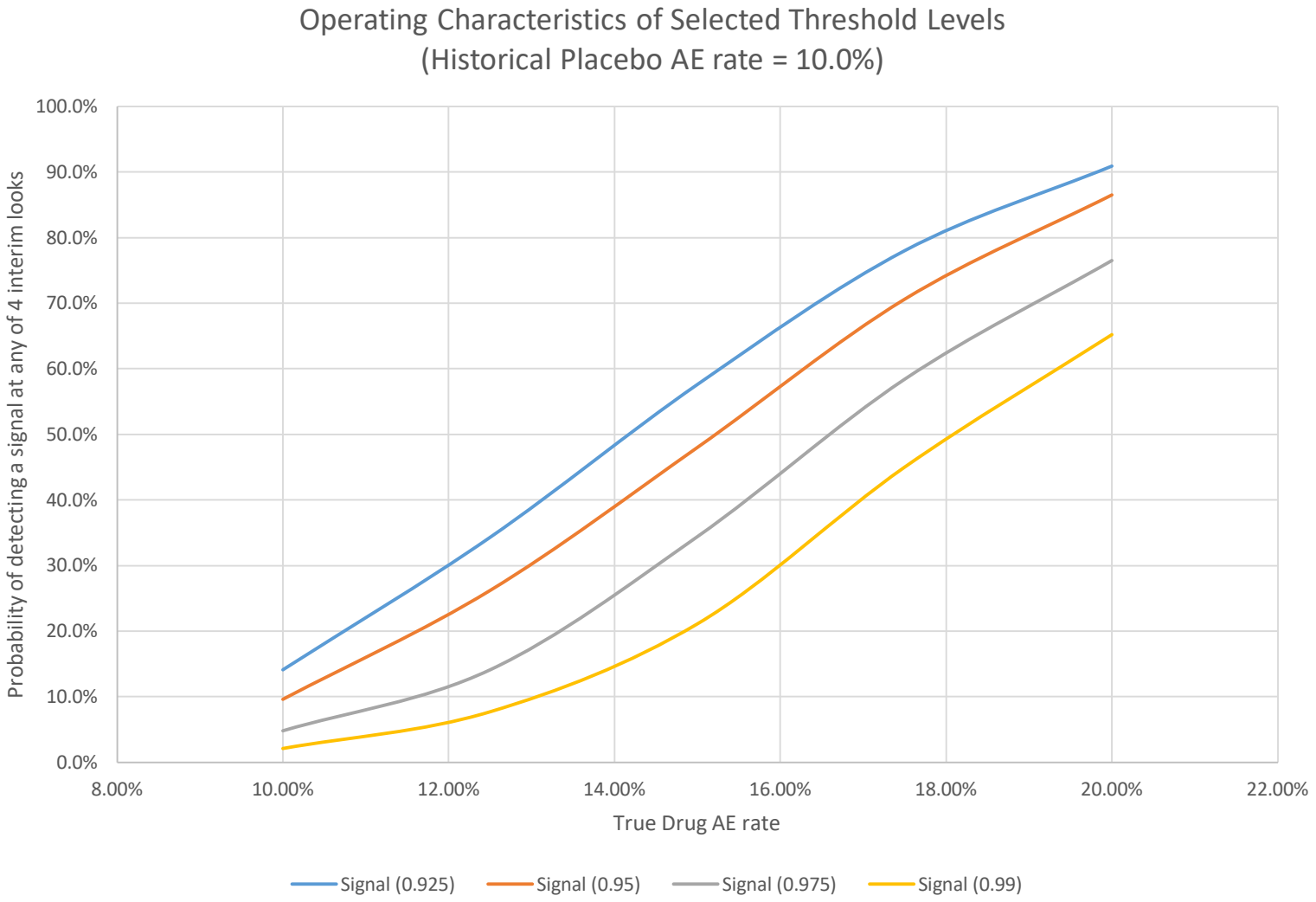
Simulation Results (comparing threshold cut-points)

The false positive rate is nearly 10% for the 0.925 threshold level, therefore probably not a good threshold to use.



Simulation Results (comparing threshold cut-points)

Recommend using 0.975 probability threshold levels, as the false positive rate for the 0.925 level is 15%, and 10% for the 0.95 level.



SSD Visualization: (Forest Plot of Frequent AEs)

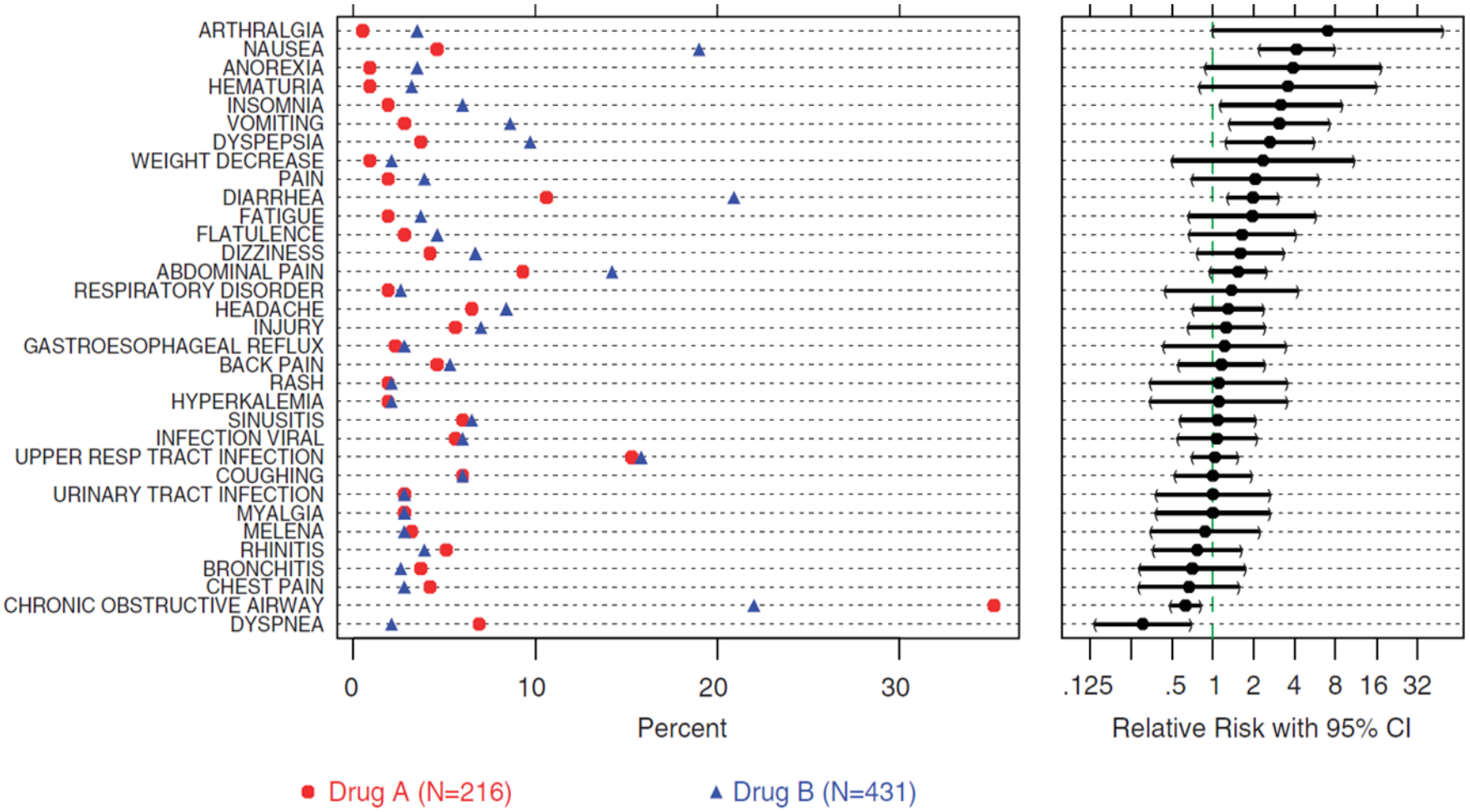
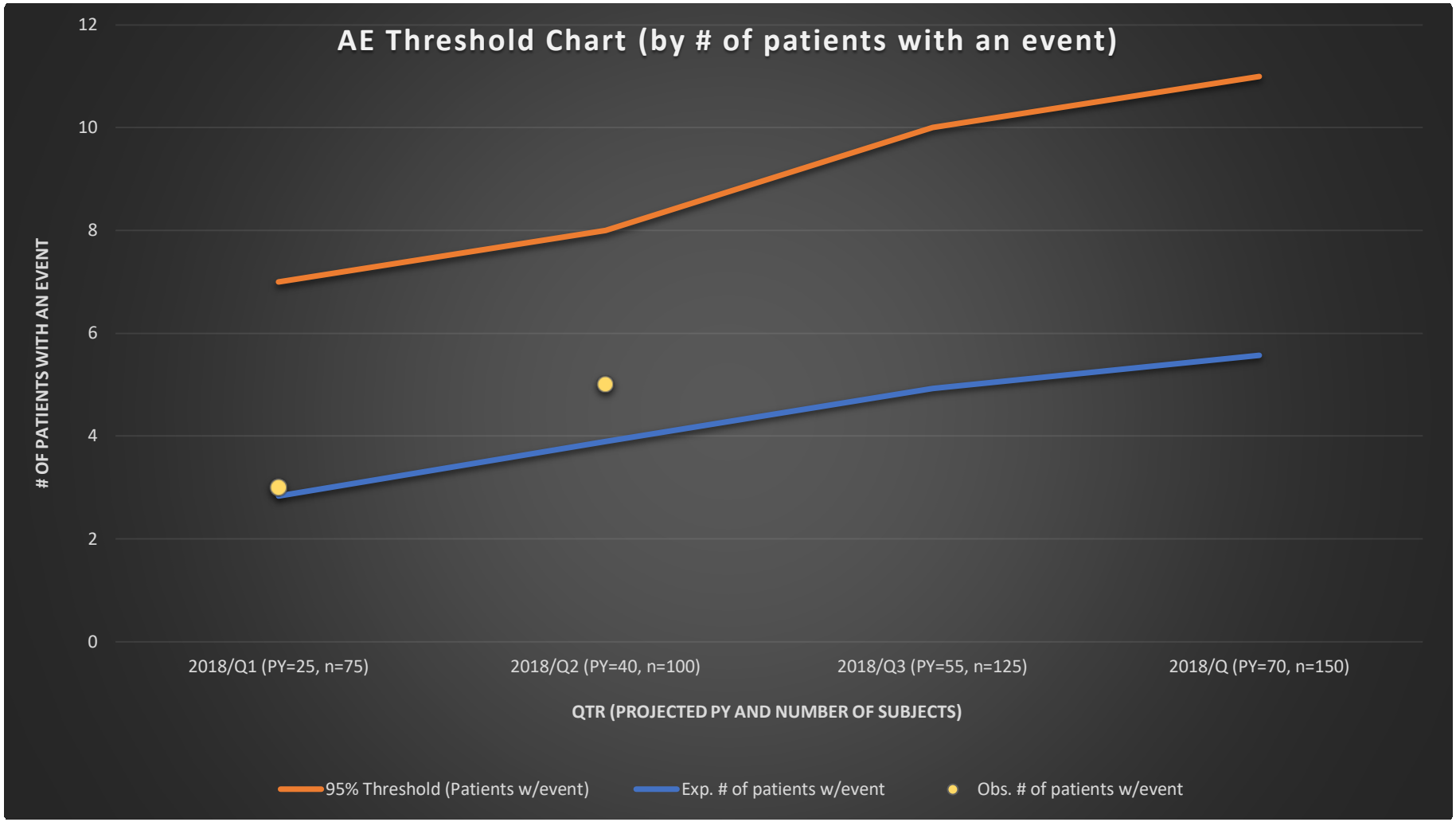


Figure 9. Most frequent on-therapy adverse events sorted by relative risk.

Source: [Amit (2008); Pharmaceutical Statistics]

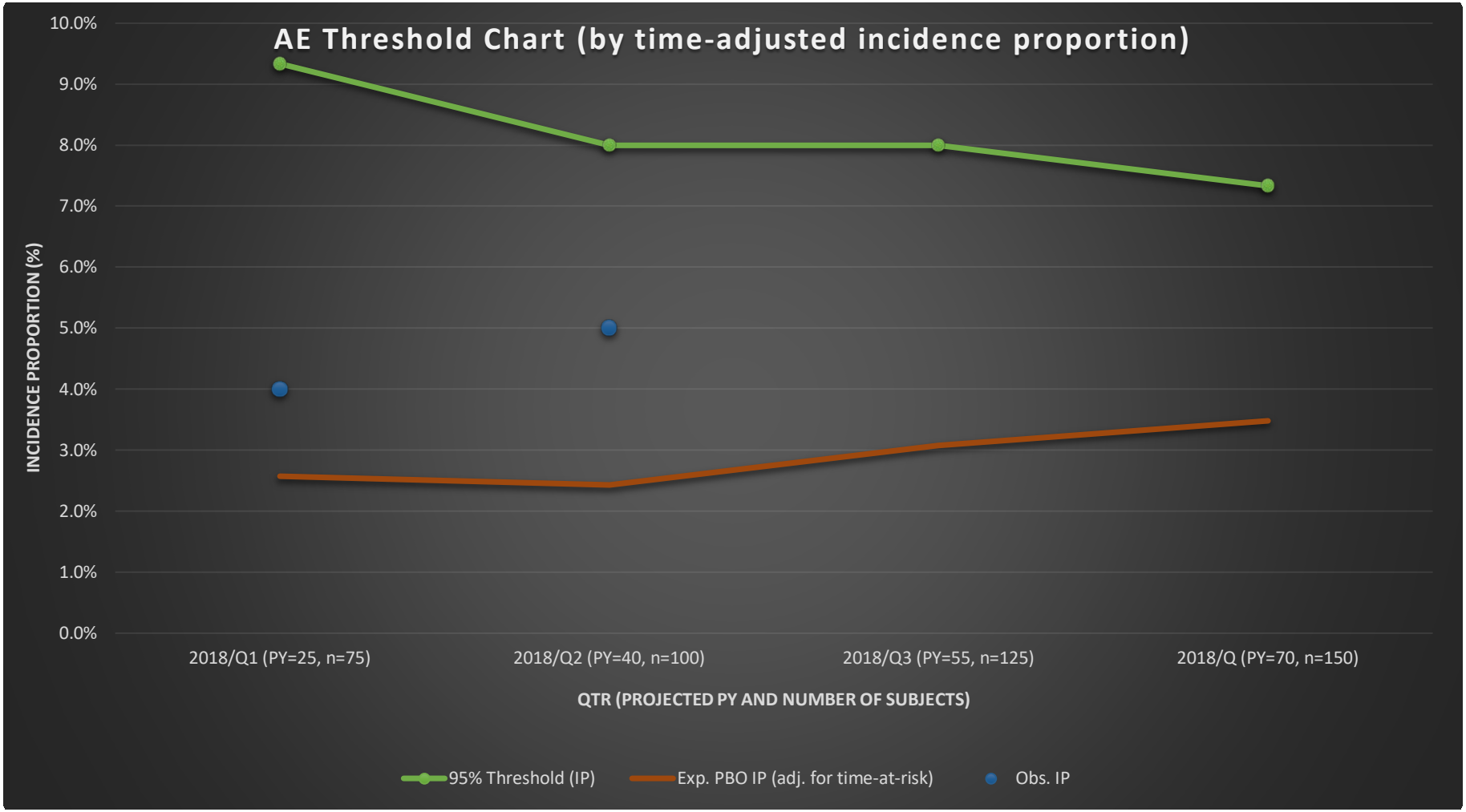
SSD visualization: (Threshold Plots)

- Incidence (count): Alternative Look



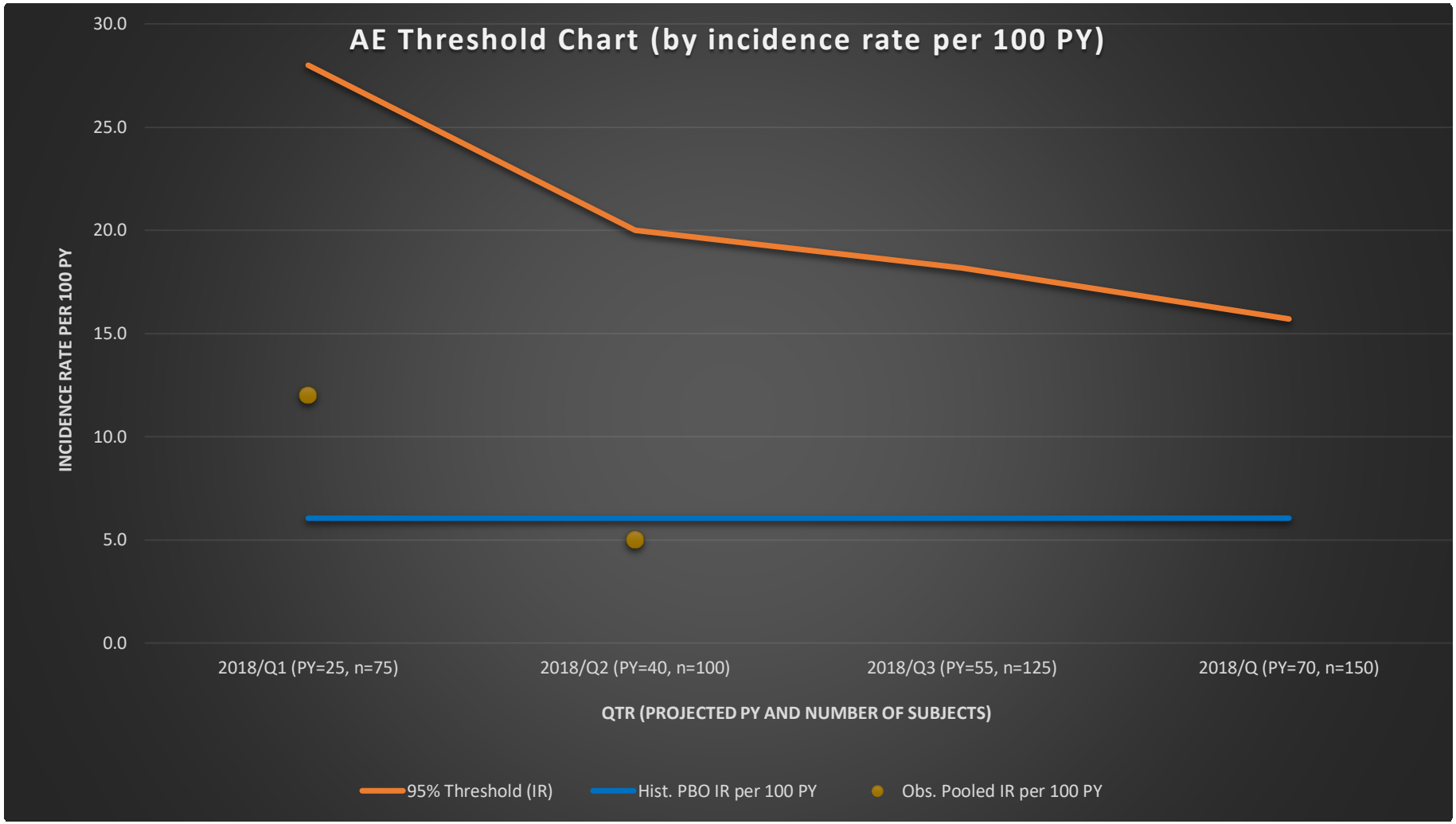
SSD visualization: (Threshold Plots)

- Incidence Proportion



SSD visualization: (Threshold Plots)

- Incidence Rate



SSD and IND safety reporting

- SSD Framework

  - SSD for AEs / Serious and Unexpected Suspected Adverse Reactions (SUSARs) [CRF 312.32(c)(1)(i)(C) and CRF 312.32(c)(2)]

    - Assessment of all adverse events
    - Specific assessment for SUSARs (including life threatening and fatal) for IND safety reporting needs

  - SSD for Serious (Expected) Suspected Adverse Reactions [CRF 312.32(c)(1)(iv)]

    - Increased rates of occurrence over that listed in the protocol or IB

SSD and IND safety reporting

- SSD for SUSAR Reporting
  - Phase 2-4 trials
  - 1st relevant randomized trial (e.g. little or no information for drug from randomized longitudinal trial)
    - Use model which require no prior for drug group, or use weak prior (e.g. low influence)
  - One or more relevant randomized trials available for the compound
    - Use information from the completed relevant trial(s) to inform priors for the drug group (and placebo group)
  - Threshold setting: use information from prior relevant trials (from a historical compound for the same disease indication, with the same or very similar patient population).
  - Confirmation of signal:
    - For events that exceed thresholds (from blinded analysis) and meet other important medical criteria, to assess and confirm causality, the SAC (or relevant body) can conduct additional Bayesian analysis on unblinded data, and include information from all other relevant trials, as well as information from the historical placebo database (to inform priors).

SSD and IND safety reporting

➢ SSD for Serious (Expected) Suspected Adverse Reactions Reporting

- ❑ Phase 2-4 trials
- ❑ Objective: Looking for increased rates of occurrence over that listed in the protocol or IB
- ❑ Use a model (e.g. population mixture models) which incorporates the prior relevant completed randomized trials (to inform priors for drug and placebo groups)
  - Modify analysis, so that historical comparisons and thresholds are based upon estimates for drug treatment groups, instead of historical placebo.
- ❑ Threshold setting: use information from prior relevant trials (for the current compound under investigation – drug treated patients).
- ❑ Confirmation of signal:
  - If evidence indicates an increased rate of serious adverse reaction for the particular event in question, then information for the event/reaction in question should be submitted to the SAC or relevant DMC.
  - An additional Bayesian analysis can be conducted in which information from all previous relevant randomized, completed trials, and unblinded data from the current ongoing trial(s) can be used to assess the degree of change in rate. The completed relevant randomized trials will be used to create informative priors.
  - New estimates for risk difference or relative risk, and corresponding posterior probability will be computed to help confirm if there is an increase in the previously reported adverse reaction rate.

abbvie

Lessons Learned / Final Thoughts / Conclusion

Additional Research

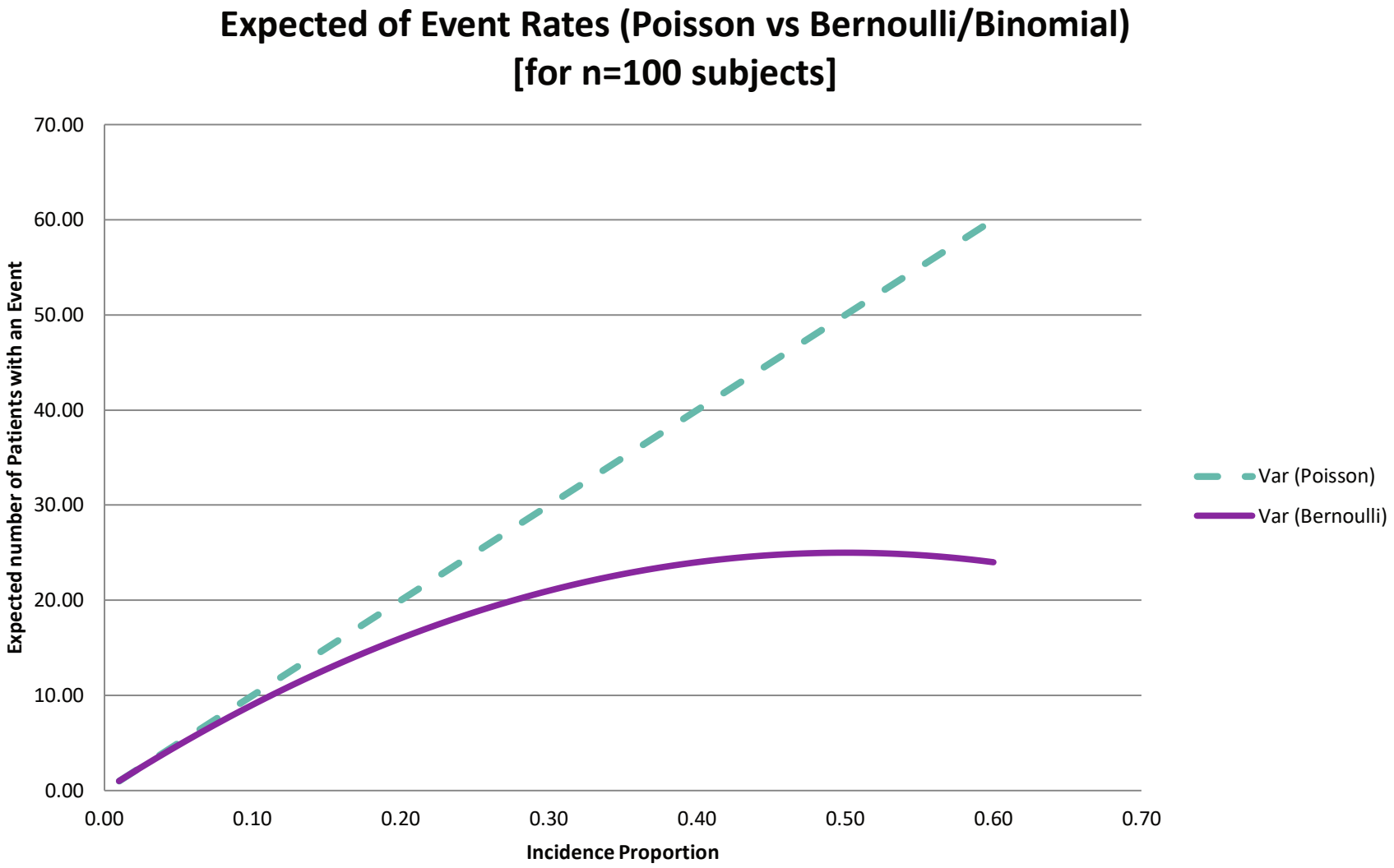
- Evaluation of Additional Models to Better Account for Variable Follow-up Time and Higher Incidence Proportions/Rates

  - For an interim look, patients will have varying amount of follow-up anywhere from a few days to several weeks or months. How does this affect the distribution of data and subsequent operating characteristics of models?

  - Do models that work for adverse events that are rare or uncommon behave accordingly for more common events or combined events (with higher incidence proportions)?

  - How to select and establish priors and how do they impact operating characteristics of models?

Additional Research



Additional Research

- Models to account for variable follow-up time (i.e. interim cut, patient withdrawals, long-term trials)
  - CDF for time to event (Exponential Distribution), assumes constant HR:

$$F(t_i) = 1 - e^{-\lambda t_i}$$

where: t_i represents time of the i^{th} patient

  - The probability that a patient experiences an event can be expressed as:

$$P(y_i=1) = 1 - e^{-\lambda t_i} = \pi_i$$

  - Therefore y_i has a Bernoulli distribution, with pmf given as:

$$p(y_i) = \pi_i^{\,y_i} (1 - \pi_i)^{\,1 - y_i}, \quad y_i = 0, 1$$

Additional Research

- Poisson-Binomial Distribution

Let Y = $\sum_{i=1}^{n} y_i$, where y_i have Bernoulli distributions with different probabilities

(e.g. π_i), due to different follow-up times, then Y has a Poisson-Binomial Distribution:

$$P[Y = y] = \sum_{A \in FY} \prod_{i \in A} \pi_i \prod_{j \in Ac} (1 - \pi_j)$$

If all patient follow-up time is identical, then (θ) = E[Y] / n is analogous to the parameter (π) from the binomial distribution. However, if patient follow-up-time is not identical, then the Poisson-Binomial Distribution (or some approximation) may be required for modeling.

Additional Research

➢ Additional Bayesian Models to Research and Consider

❑ Individual Bernoulli Likelihood Models

- (1) Population Mixture (prior) - Individual Bernoulli likelihood
- (2) Individual Mixture (prior) - Individual Bernoulli Likelihood

❑ Incidence Rate Models

- (3) Population Mixture (prior) - Poisson Likelihood

❑ Time to Event Models

- (4) Population Mixture (prior) - Exponential Likelihood

❑ Poisson-Binomial (or Approximate) Models

- Models attempting to model the Poisson-Binomial distribution
- Models approximating the Poisson-Binomial distribution [e.g. Negative Binomial Model (under-dispersed)]

Expansion of SSD

- Expansion within and beyond AE Analysis

  - Incorporation of additional features to better control historical populations (e.g. Poisson regression, propensity scores, methods/process for down-weighting historical data, etc.)
  - Develop methods for assessing non-constant hazard rates (e.g. Weibull, Gamma, piece-wise exponential, double exponential, etc.)
  - Multiplicity control (finding the proper balance)
  - Expansion to Vital signs, clinical labs, and ECG data
  - Visualization Tools and Dashboard Displays

Lessons learned / final thoughts / Conclusions

- Issues (logistical and other) to Consider
  - Getting all required safety data in a timely fashion (e.g. events and associated exposure data), to keep with real-time reporting of serious adverse events to company safety departments.
  - AE terminology
  - Synchronizing SAE reporting for IND-SR purposes, between early phase and late phase.
  - Incorporation of open label extension studies
  - Set-up of proper firewalls (e.g. SAC)
  - Incorporation of blinded SSD (and unblinded aggregate reporting) within the entire pharmacovigilance spectrum
  - Coordination with DSMBs
  - Training and understanding

Lessons learned / Final thoughts / Conclusions

- Incidence Proportion or Incidence Rate?

  - Incidence Proportion (Time-adjusted) Issues:
    - Computation of expected incidence (count) and corresponding time-adjusted incidence proportion is more complicated
    - Incorporating exposure time (e.g. follow-up time and time-at-risk) can be tricky
    - Average time-at-risk per patient versus complete distribution of time at risk for (across all patients)
    - Binomial/Bernoulli likelihood model for ongoing trials (with various rates [θ_i]due to different follow-up times)
  - Incidence Rate Issues:
    - clinicians are more accustomed to seeing and understanding incidence proportions, not as comfortable with incidence rates
    - Regulators are accustomed to incidence proportion and labels report incidence proportion, not incidence rate
    - An actual count of patient with the event is easier to digest for SSD, then the more abstract incidence rate
    - Beta-binomial distribution (within Bayesian framework, is easy to work with)

Lessons learned / final thoughts / conclusions

- Final Thoughts / Summary
  - Continued development of SSD and all Standardized Safety Evaluation and Analysis is Critical
  - The Bayesian Framework provides a useful tool for conducting SSD analysis with continuous updating and aggregation of clinical trials data
  - Simulations provide a comprehensive method for evaluating methods and testing ideas
  - The models evaluated produce approximately comparable results in most cases (e.g. “all roads lead to Rome”):
  - Expansion to other safety domains should be developed:
    - Clinically significant lab and ECG findings can be analyzed using a similar Bayesian framework
    - Models for vital signs (e.g. BP, HR) should be considered – continuous data framework

Questions?

Thanks

References

Amit, O. Heiberger, R. Lane, P. (2008). Graphical Approaches to the Analysis of Safety Data from Clinical Trials. *Pharmaceutical Statistics*; 7(1): 20-35.

Crowe, B. et al. (2009). Recommendations for Safety Planning, Data Collection, Evaluation and Reporting During Drug, Biologic and Vaccine Development: *A Report of the Safety Planning, Evaluation and Reporting Team.* Clinical Trials; 6: 430-440.

Crowe, B. Chuang-Stein, C. Lettis S. Brueckner, A. (2016). Reporting Adverse Drug Reactions in Product Labels. *Therapeutic Innovation & Regulatory Science*; 50(4): 455-463.

FDA CDER CBER (2005). Guidance for Industry: Good Pharmacovigilance Practices and Pharmacoepidemiological Assessment. Food and Drug Administration, Silver Spring, MD.

FDA CDER CBER (2012). Guidance for Industry and Investigators: Safety Reporting Requirements for INDs and BA/BE Studies. Food and Drug Administration, Silver Spring, MD.

FDA CDER CBER (2105). Draft Guidance for Industry: Safety Assessment for IND Safety Reporting. Food and Drug Administration, Silver Spring, MD.

Investigational New Drug Application (2018). *Code of Federal Regulations*. Title 21 Volume 5. Section 312.32 IND Safety Reporting. [21CFR312.32].

Kerman, J. (2011). Neutral Noninformative and Informative Conjugate Beta and Gamma Prior Distributions. Electronic Journal of Statistics; 5: 1450-1470. DOI: 10.1214/11-EJS648.

References (cont.)

Report of the Council for International Organization of Medical Sciences (CIOMS) Working Group VIII, 2010, Practical Aspects of Signal Detection in Pharmacovigilance. Report of CIOMS Working Group VIII.

Wen, S. Ball G. Dey, J. (2015). Bayesian Monitoring of Safety Signals in Blinded Clinical Trial Data. Annals of Public Health and Research; 2(2): 1019.

Yao, B., Zhu, L., Jiang, Q., Xia, A. (2013). Safety Monitoring in Clinical Trials. *Pharmaceutics;* 5: 94-106. doi: 10.3390/pharmaceutics5010094


abbvie

abbvie

Back-up Slides

Proposed Bayesian Models: New Model Specifications

- Common parameters, random variables, and fixed constants for all models

λ_{HP} = underlying incidence (hazard) rate from a historical placebo population (i.e. referred to as λ_w in earlier slides)

PT_{HP} = total patient time-at-risk from a historical placebo population

Y_i = A random variable that takes on the value of 1 (if patient i has the respective adverse event) or 0 (if patient does not have the adverse event), for a new blinded study or set of studies.

RR = randomization ratio

Proposed Bayesian Models: New Model Specifications

- (1) Population Mixture (Prior) – Individual Bernoulli Likelihood Model

$Y_i \mid trt, \lambda_p, \lambda_d \sim Bernoulli(\pi_i)$; $i=1,\ldots,N$

where: $\pi_i = \pi_{pi}(trt) + \pi_{di}(1-trt)$

$\pi_{pi} = 1 - e^{-\lambda pti}$; $i = 1,\ldots,N$

$\pi_{di} = 1 - e^{-\lambda dti}$; $i = 1,\ldots,N$

$trt \sim Beta(\alpha,\beta)$; where $\alpha+\beta$ = reasonable ESS (e.g. 800), and $\beta=(RR)(\alpha)$

$\lambda_p \sim Gamma(\alpha_p = \lambda_{HP} \times PT_{HP}, \beta_p = PT_{HP})$*

$\lambda_d \sim Gamma(\alpha_d, \beta_d)$ **

* (Modify accordingly to incorporate placebo information from completed trials of current compound.)

** *(α_d, β_d can be set to values [e.g. α_P/x and β_p/x] or some other value, so that λ_d is a neutral prior [e.g. x takes on value so that β/x approximates PT for 1 subject in the study], when no prior knowledge of drug is available; or set to relevant values elicited from previously completed studies of the current compound.)*

Proposed Bayesian Models: New Model Specifications

- (2) Individual Mixture (Prior) – Individual Bernoulli Likelihood Model

$$Y_i \mid trt_i, \lambda_p, \lambda_d \sim \text{Bernoulli}(\pi_i) \;\; ; \;\; i=1,\ldots,N$$

$$\text{where: } \pi_i = \begin{cases} 1 - e^{-\lambda pti}, & \text{if } trt_i = 1 \text{ (e.g. pbo group)} \\ 1 - e^{-\lambda dti}, & \text{if } trt_i = 0 \text{ (e.g. drug group)} \end{cases}$$

$trt_i \sim \text{Bern}(pla) \;\; ; \; i = 1,\ldots,N$

$pla \sim \text{Beta}(\alpha,\beta)$; where $\alpha+\beta$ = reasonable ESS (e.g. 800), and $\beta=(RR)(\alpha)$

$\lambda_p \sim \text{Gamma}(\alpha_p = \lambda_{HP} \times PT_{HP}, \beta_p = PT_{HP})$*

$\lambda_d \sim \text{Gamma}(\alpha_d, \beta_d)$ **

* (Modify accordingly to incorporate placebo information from completed trials of current compound.)

** *(α_d, β_d can be set to values [e.g. α_p/x and β_p/x] or some other value, so that λ_d is a neutral prior [e.g. x takes on value so that β/x approximates PT for 1 subject in the study], when no prior knowledge of drug is available; or set to relevant values elicited from previously completed studies of the current compound.)*

Note: This model takes an extremely long time to run, with SAS Proc MCMC

Proposed Bayesian Models: New Model Specifications

- (3) Population Mixture (Prior) – Poisson Likelihood Model

$Y \mid trt, \lambda_p, \lambda_d \sim Poisson(\lambda T)$

where: $\lambda = (\lambda_p) \times (TRT) + (\lambda_d) \times (1 - TRT)$

T = total time-at-risk

$trt \sim Beta(\alpha,\beta)$; where $\alpha+\beta$ = reasonable ESS (e.g. 800), and $\beta=(RR)(\alpha)$

$\lambda_p \sim Gamma(\alpha_p = \lambda_{HP} \times PT_{HP}, \beta_p = PT_{HP})$

$\lambda_d \sim Gamma(\alpha_d, \beta_d)$ **

Note: The Poisson distribution may not be the best choice with larger incidence rates, as the distribution (i.e. variance) of the expected counts is likely to be under-dispersed.

Proposed Bayesian Models: New Model Specifications

- (4) Overall Population Mixture (Prior) – Individual Exponential Likelihood Model (Time-to-event)

$r,T \mid \beta_0 \sim$ Exponential(λ) ; in which the log of the likelihood function is used

The log-likelihood (LL) function is derived as: $r \ln(\lambda) - \lambda T$

where: r = number of patients with the event,

T = total time for all patients (to event or censor)

Link any covariates to λ, with $\lambda i = \exp(x_i` \beta)$;

for overall model: $\beta_0 = \ln(\lambda)$, $\lambda = e^{\beta_0}$

then $LL = r \beta_0 - T e^{\beta_0}$

$\beta_0 \sim$ Normal(0, 10,000) ; diffuse prior

$\lambda_p \sim$ Gamma($\alpha_p = \lambda_{HP} \times PT_{HP}$, $\beta_p = PT_{HP}$)*

* For comparison purpose (e.g. $\lambda > \lambda_p$)

Note: Modify β accordingly to incorporate individual treatment group (latent variable) priors