Bayesian detection of potential safety signal from blinded clinical trial data

Saurabh Mukhopadhyay, Ph.D
Brian Waterhouse, M.Sc.

DIA Bayesian Scientific Working Group KOL Lecture Series
June 2021

abbvie

Data & Statistical Sciences

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

Saurabh Mukhopadhyay is an employee of AbbVie Inc. and may own AbbVie stock. He can be contacted at saurabh.mukhopadhyay@abbvie.com

Brian Waterhouse is a former employee of AbbVie Inc. and may own AbbVie stock. He can be contacted at brian.waterhouse@merck.com

Outline

Part I: Methods

Part II: Implementation

Part I: Methodologies

- Key Considerations
- Motivations and Challenges for Blinded Signal Detection
- Exposure Based Modeling
- BDRIBS framework
- Posterior Inference to Screen Signal
- Predictive Inference
- Determining Decision Thresholds using BF
- Incorporating Uncertainty of Historical Information
- Prospective Planning using Contour Plots
- Concluding Remarks

Key Considerations

- Ensuring patients' safety in clinical trials is extremely important
  - Major responsibility of the sponsors of the clinical trials
  - Regulatory authorities encourage aggregate review of safety from ongoing trials
  - Specifically to identify potential risks of adverse events of special interest

- FDA recommends[1] and discusses that sponsors develop
  - Safety Assessment Committee (SAC): to review important safety information at appropriate intervals to make a recommendation whether an event or group of events meets the criteria for IND safety reporting
  - Safety Surveillance Plan (SSP): to prospectively plan the processes and procedures for assessing serious adverse events (SAEs) and other important safety information

[1] Safety Assessment for IND Safety Reporting, FDA Guidance for Industry, Dec 2015

Motivations and Challenges for Blinded Signal Detection

- Sponsors may leverage an appropriate blinded review of safety data
  - To meet these needs
  - To ensure clinical trial integrity
- FDA suggested that an internal SAC conduct unblinded aggregate reviews
  - Routine unblinded reviews can create internal firewall issues, namely, unintentional unblinding or the appearance of inappropriate unblinding of study personnel
  - Blinded monitoring can alleviate those problems and help to ensure trial integrity

- Challenges for blinded monitoring
  - Methodological challenges on how to make ongoing inference from blinded combined data
  - Operationally can be a complex endeavor as well
    - Will require a prospective planning of all appropriate steps of how and when
    - Must be a multi-disciplinary approach

Exposure-based modeling

- For drug safety monitoring, exposure-time is extremely important
  - It would allow consideration of differential exposure of patients
  - It would also allow adjustment for exposure when combining studies that did not start at the same time - combined monitoring (blinded or not) is critical in safety monitoring programs

- Modeling events over time
  - Let δ denote the rate at which the events occur per unit exposure time (say, in 1 patient-year)
  - Let E_t and Y_t denote the total exposure time and the number of events of interest (EOI) at time t
  - Since incidence rate for an SAE is typically very small and large patient-years are often at risk – we assume that number of events follow a Poisson process over time
  - The Poisson distribution describing this process is thus

$$P[Y_t = k | E_t = \tau, \delta] = e^{-\delta \cdot \tau} \frac{(\delta \cdot \tau)^k}{k!}, k = 0, 1, 2, \ldots$$

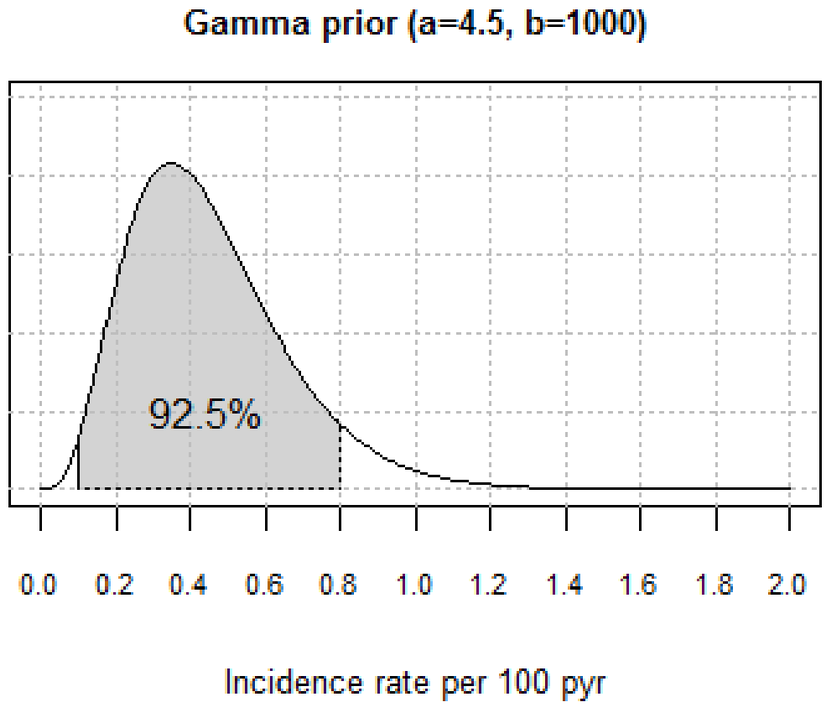
Assigning prior on incidence rate δ

- We assume a prior distribution for δ:

$$\pi(\delta)$$

  - Historical knowledge on the concerned SAE can be used to model $\pi(\delta)$

- Example: Suppose we know from the past data that the average rate was about 4.5 per 1000 patient-years (pyr) and the chance of observing more than 1 events in 100 pyr is very low (say < 2.5%)
  - Gamma distribution is often used to model the prior since it is a natural conjugate to Poisson distribution
  - This information may be modelled well using a $Gamma(a, b)$, with a =4.5 and b=1000. It gives a prior mean of 4.5 per 100 pyr and the chance of observing more than 1 events in 100 patient years is about 2%.



BDRIBS Framework

- BDRIBS (Bayesian Detection of potential Risk using Inference on Blinded Safety data)
  - A complete Bayesian framework to monitor and detect potential risk of safety signal from ongoing blinded clinical trial data
- Consider an ongoing blinded clinical trial:
  - Suppose we observe (Y, E) where Y is the total number of events and E is the total exposure (in patient-years)
    - Subjects were randomly allocated with k: 1 ratio to treatment group (T) and control group (C).
    - Let the true (unknown) incidence rates be δ_0 and δ_1 for control and treatment group, respectively
    - To monitor the safety signal by studying the relative risk $\mathrm{r} = \delta_1/\delta_0$
- Note that the relative risk $\boldsymbol{r}$ is the **parameter of interest**

Distribution Assumptions and Derivation

- Although unobservable from blinded data, let (Y_0, E_0) and (Y_1, E_1) be the pairs of event and exposure time in C (index=0) and T (index=1) groups, respectively

  We denote true mean number of events as λ_0 and λ_1, thus

  $$\lambda_j = \delta_j * E_j \quad (j = 0,1)$$

  Number of events in each group, Y_j, are independent Poisson:

  $$Y_j \sim Poisson(\lambda_j)\ j = 0,1$$

  Denoting $Y = Y_0 + Y_1$ and we have $Y \sim Poisson(\lambda = \lambda_0 + \lambda_1)$

  Since, $\delta_1 = \mathrm{r} \cdot \delta_0$ we get $\lambda_0 + \lambda_1 = \delta_0 \cdot (E_0 + r \cdot E_1)$

  Assuming similar dropout rates in both groups, the total exposure E would be distributed approximately in the same proportion as in the allocation ratio $(k : 1)$, thus $E_1 \approx k \cdot E_0$

$$Y \sim Poisson\left(\lambda = E \cdot \delta_0 \cdot \frac{r \cdot k + 1}{k + 1}\right)$$

Prior modeling of the risk parameter r

- The likelihood function of observed blinded data $(Y = y, E = e)$ is thus a Poisson pmf which we denote as

$$f(y|e, k, r, \delta_0) \equiv f(y|r, \delta_0)$$

Assuming δ_0 will be estimated from historical data and denoting $\pi(r|\delta_0)$ as the prior of r for a given δ_0 the corresponding posterior is thus:

$$\pi^*(r|y, \delta_0) \propto f(y|r, \delta_0) \cdot \pi(r|\delta_0)$$

We typically do not have any information on the relative risk r – so we need to assign a noninformative prior on r

Note that it would be harder to define a non-informative prior on r directly as it is defined on [0,∞).

Using theoretical result that the conditional Poisson is a Binomial we have $Y_1|Y \sim Bin(Y, p)$ where $p = \frac{\lambda_1}{\lambda_1+\lambda_2}$ and with simplification

$$p = \frac{k \cdot r}{k \cdot r + 1} \quad or\ r = \frac{p}{k \cdot (1-p)}$$

where p = Prob(an event has occurred in the treatment group | given that the event occurred in the combined blinded data)

This interpretation of p helps us to assign an appropriate [noninformative] prior on p.

Then we use the mathematical relationship between p and r to induce the prior on *r*

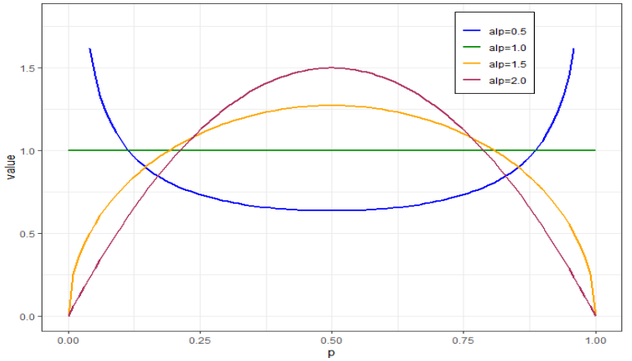
Prior modeling of p and r

- For equal allocation, i.e., when $k = 1$, if we have no prior knowledge of additional risk for the treatment group as compared to the control group, then it is appropriate to assign a non-informative prior on p
- Jeffreys' non-informative prior, Beta(0.5,0.5), or the Uniform prior, U(0, 1), may be reasonable choices
- We can then obtain the corresponding prior densities of r – we notice that this gives equal prior weight on r < 1 and r >1:

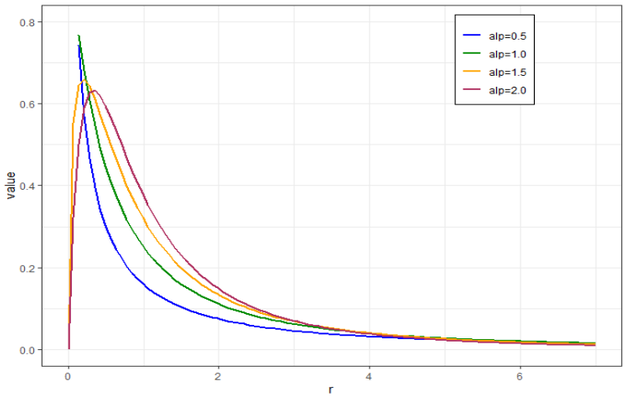
$$P(r > 1) = P(r < 1) = 0.5$$

- This holds for any symmetric prior on p, e.g., $p \sim Beta(alp, alp)$ but with larger values of alp we would make it harder to detect increased signal

Priors for $p \sim Beta(\alpha, \alpha)$



Induced prior densities for r



Prior modeling of r (when $k > 1$)

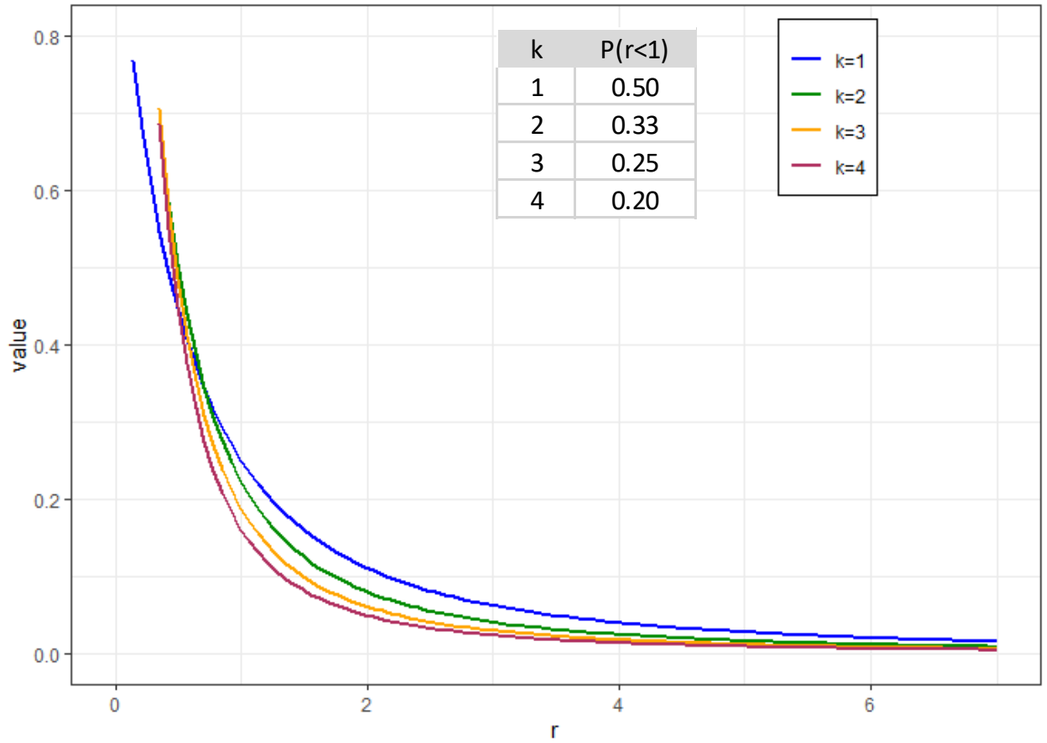
- If we keep using symmetric priors such as Beta(0.5,0.5), or U(0,1) when $k > 1$, then it puts decreasing weight on $r > 1$ as k gets larger
- When $p \sim U(0,1)$ then

$$P(r > 1) = 1/(k+1)$$

- There are a few ways we can adjust for this
  1. Use prior of r derived from $k = 1$.
  2. Adjust prior on $p \sim beta(alp, bet)$ so that $P(r > 1)$ remains about 0.5
- Note that Bayes Factor can counterbalance some of this without any upfront adjustment in prior

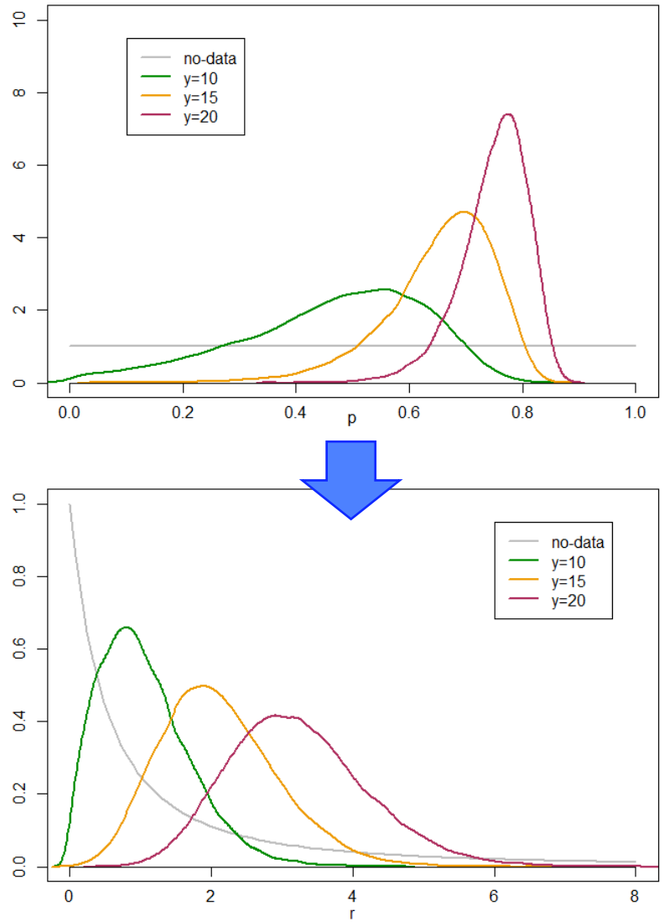
Induced prior densities for r when $p \sim U(0,1)$

k	P(r<1)
1	0.50
2	0.33
3	0.25
4	0.20



Posterior inference using BDRIBS – A worked example

- Suppose that with a total of $E = 2000$ patient years at risk with equal allocation $(k = 1)$ we see a total of y events (i.e., $Y = y$)
  - We first proceed with a conditional Bayesian inference by assuming δ_0 to be fixed at $\widehat{\delta_0}$ estimated from past historical data:
    - Suppose that from past historical data we know that $x = 18$ events had occurred in the control group from H=4000 patient-years of exposure resulting in an incidence rate estimate of $\widehat{\delta_0} = \frac{x}{H} = 0.0045$
  - Let us assume $p \sim U(0,1)$
  - Let us now consider a few what-if values $y = 10, 15, 20$ to assess the risk signals



Since prior was not conjugate to likelihood, needed MCMC simulations to generate samples from the posterior distribution.

Bayesian posterior inference

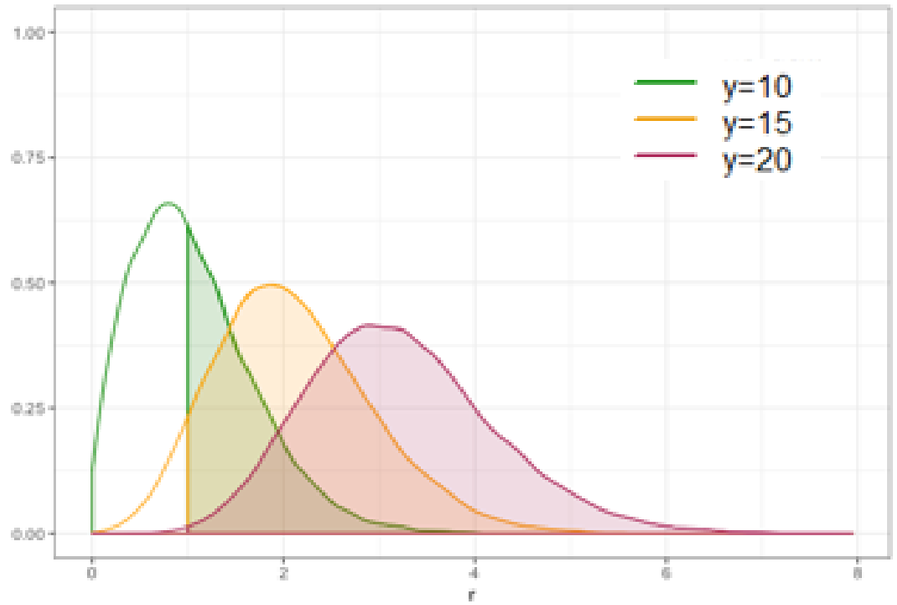
- We can compute posterior probability of $\{r > 1\}$ to assess ***any*** elevated risk

$$P(r > 1|y)$$

- On many occasions, there may be more concern if the relative risk exceeds a clinically important level c $(c > 1)$

$$P(r > c|y)$$

	P(r>1|y)	P(r>1.2|y)	P(r>1.5|y)
No-data	0.500	0.455	0.400
y=10	0.472	0.359	0.221
y=15	0.927	0.870	0.757
y=20	0.998	0.994	0.979

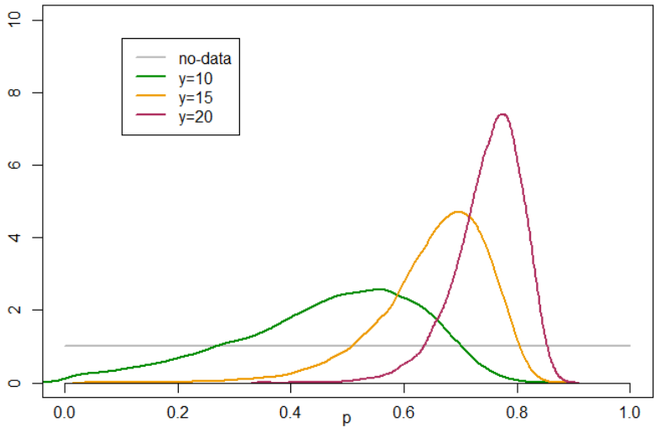


Bayesian predictive inference

- We can also leverage the framework to predict the probability that a future event will come from the treatment group
  - Recall that $Y_1 | \, Y, \mathrm{p} \sim Bin(Y, p)$
  - Thus, the predictive probability that a future event will come from the treatment group is given by

$$P(Y_1^{new} = 1 | Y^{new} = 1, y) = \int p \cdot \pi(p|y) dy = E[p|y]$$

which is simply the posterior mean of p for our model



	E[p|y]
No-data	0.500
y=10	0.465
y=15	0.654
y=20	0.750

How to do a Bayesian testing on signal strength

- In Bayesian inference, the use of Bayes Factors (BF) provides a decision theoretic framework to test say

$$H_0 : r \leq c \; vs. \; H_1 : r > c$$

(see e.g., Berger, 1985; Kass and Raftery, 1995; Goodman, 1999)

  - Bayes Factor of H_1 to H_0 is then defined as the ratio of marginal likelihoods and is often easier to compute as the ratio of posterior odds to prior odds:

$$\frac{P(H_1|y)}{P(H_0|y)} = \frac{P(H_1)}{P(H_0)} * BF$$

(posterior odds) = (prior odds) * (Bayes Factor)

- For example, when $y = 15$ the posterior odds for $r > 1$ is given as

$$\frac{P(r > 1|data)}{P(r \leq 1|data)} \approx \frac{0.927}{0.073} = 12.7$$

and the prior odds = 0.5/0.5 =1.

  - Thus BF [r>1|y=15] = 12.7

	P(r>1|y)	P(r>1.2|y)	P(r>1.5|y)	BF[r>1]	BF[r>1.2]	BF[r>1.5]
No-data	0.500	0.455	0.400			
y=10	0.472	0.359	0.221	0.9	0.7	0.4
y=15	0.927	0.870	0.757	12.7	8.0	4.7
y=20	0.998	0.994	0.979	539.5	198.4	68.4

Finding threshold using Jeffreys scale

- A value of BF > 1 means that data supports H_1 more than H_0
- There are limits on changes in a weight of evidence (i.e., a change in an odds ratio, or BF) that humans can reasonably perceive their degree of belief in a hypothesis in everyday use (Good, 1979)
- Scales are suggested by Jeffreys (1961) to interpret the BF:

BF	Strength of Evidence in favor of H_1
<1	Negative (supports H_0)
1 to 3	Barely worth mentioning
3 to 10	Substantial
10 to 30	Strong
30 to 100	Very strong
> 100	Decisive

- There is another slightly different scale suggested by Kass and Raftery (1995)
- We can compute corresponding thresholds on posterior probability scale (see Mukhopadhyay et al , 2018)

Sensitivity analyses and other prior modelling options

- Examining the sensitivity with respect to prior on p

  - The results so far were based on $p \sim U(0,1)$ and we could check with some other priors e.g., the Jeffreys prior on p

$$p \sim Beta(0.5, 0.5)$$

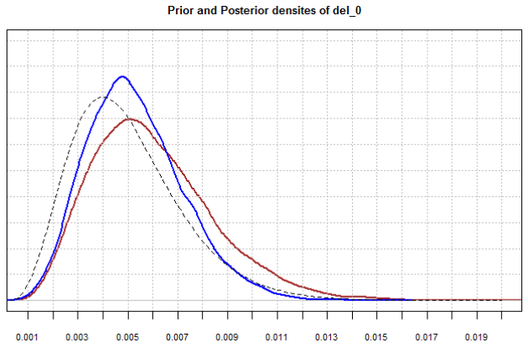
- We can also work with a joint prior modelling framework:

$$\pi(r, \delta_0) = \pi(r|\delta_0) * \pi(\delta_0)$$

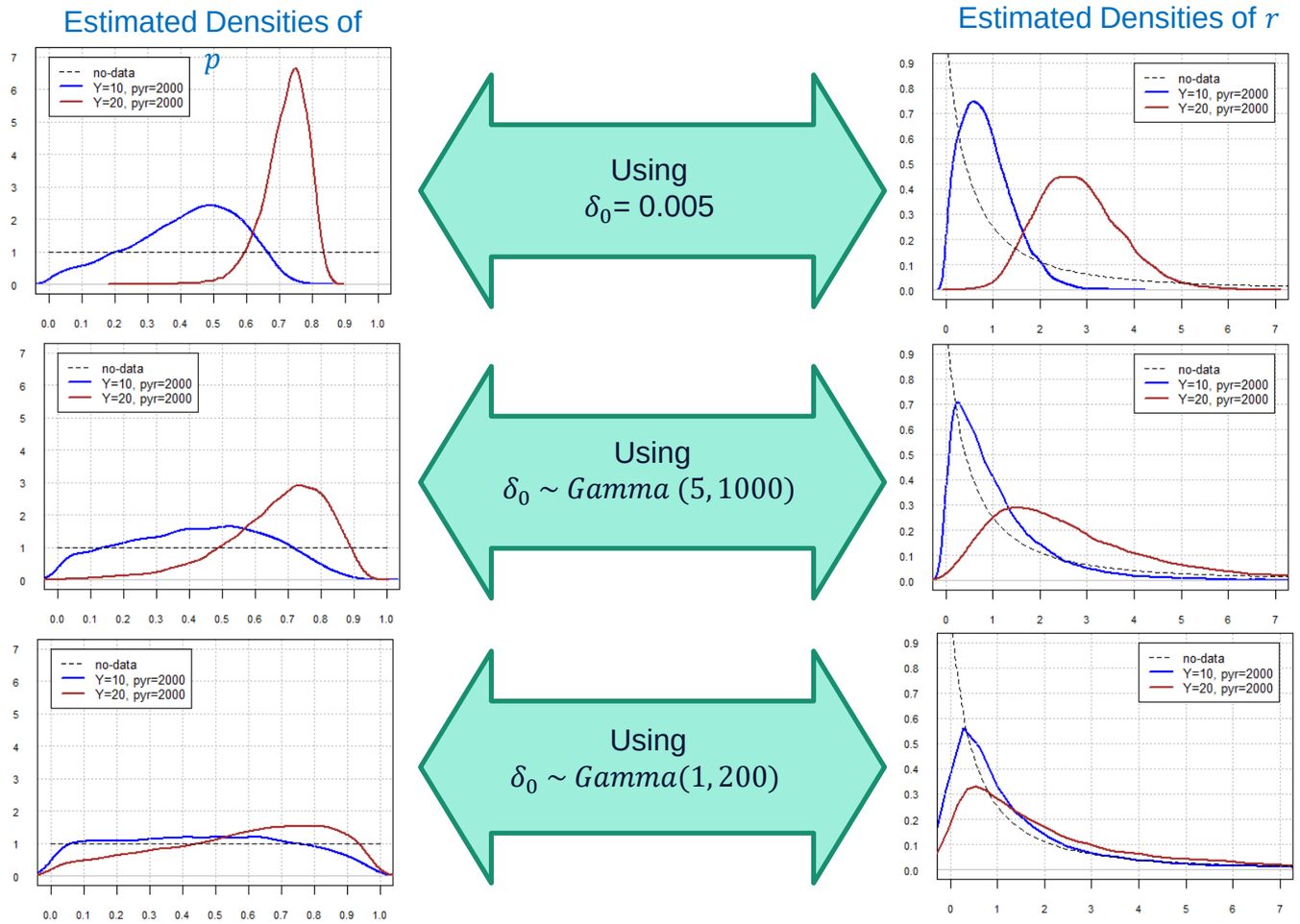
  - However, there is no gain in information on estimating δ_0 with blinded data

Average Bayesian Inference

- We can assume a prior on δ_0
  - For example, suppose $\delta_0 \sim Gamma(x, H)$ based on say, $x = 5$ historical events from $H = 1000$ patient-years in the control group
  - We do not however gain additional information on δ_0 with the new blinded data

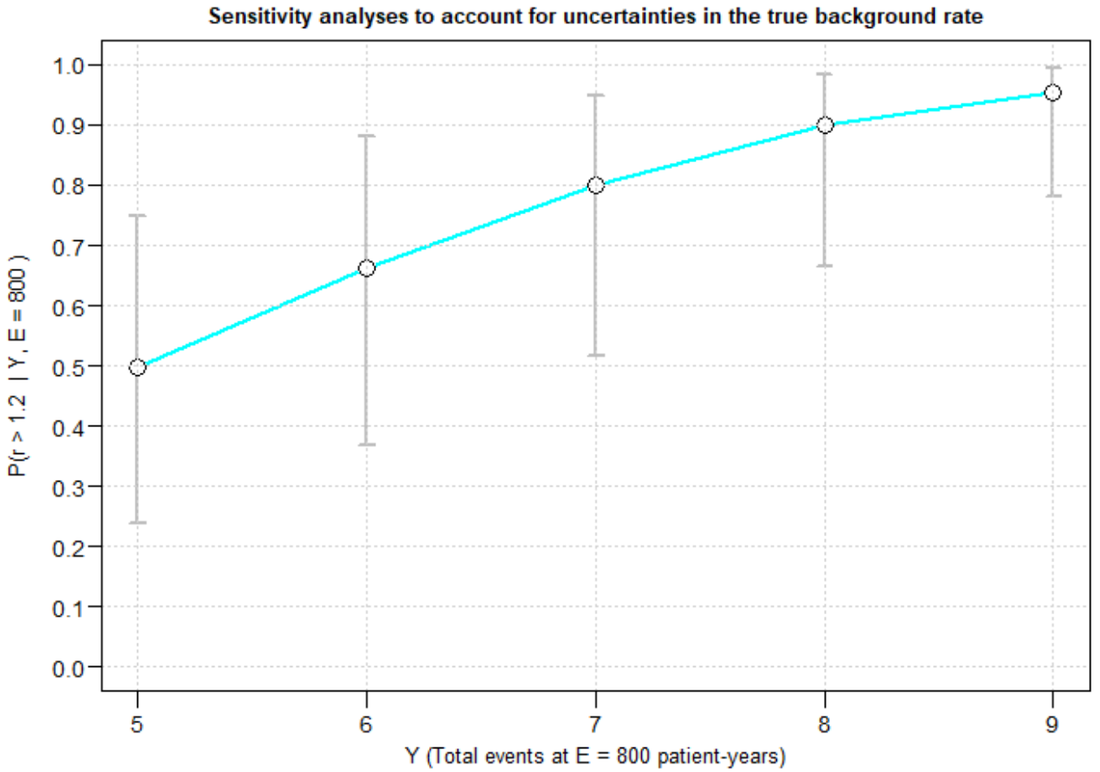


  - Thus it adds extra noise to reflect uncertainty on the background rate



Incorporating uncertainty in historical information

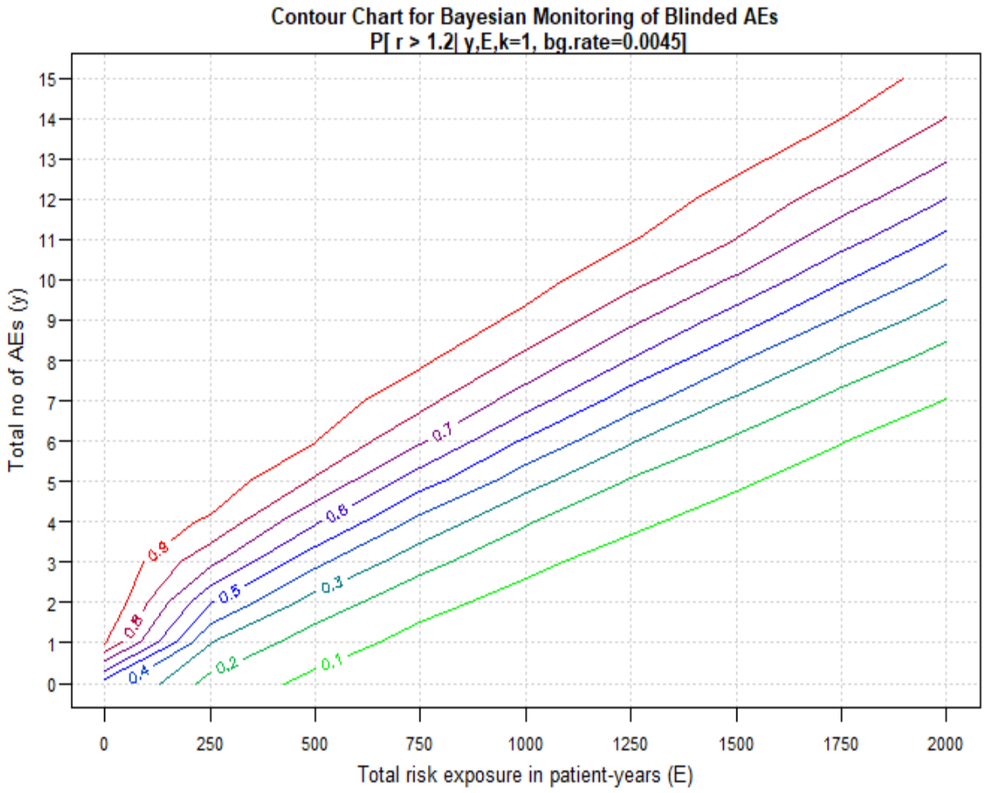
- Examining with respect to using a fixed $\widehat{\delta_0} = 0.0045$
  - We can check how the inference is sensitive with the choice we made for δ_0 by using a range of values δ_0 supported the prior based on historical data
  - For example, suppose $\delta_0 \sim Gamma(x, H)$ is based on say, $x = 18$ historical events from $H = 4000$ patient-years in the control group
  - Then the true incidence rate in the control group δ_0 would vary between 0.003 to 0.006 with ~ 90% confidence
  - We can then check the range of $P(r > 1.2|\mathrm{Y} = \mathrm{y}, \mathrm{E})$ as δ_0 would vary between 0.003 to 0.006 with mean of 0.0045



Prospective Planning using Contour Plots

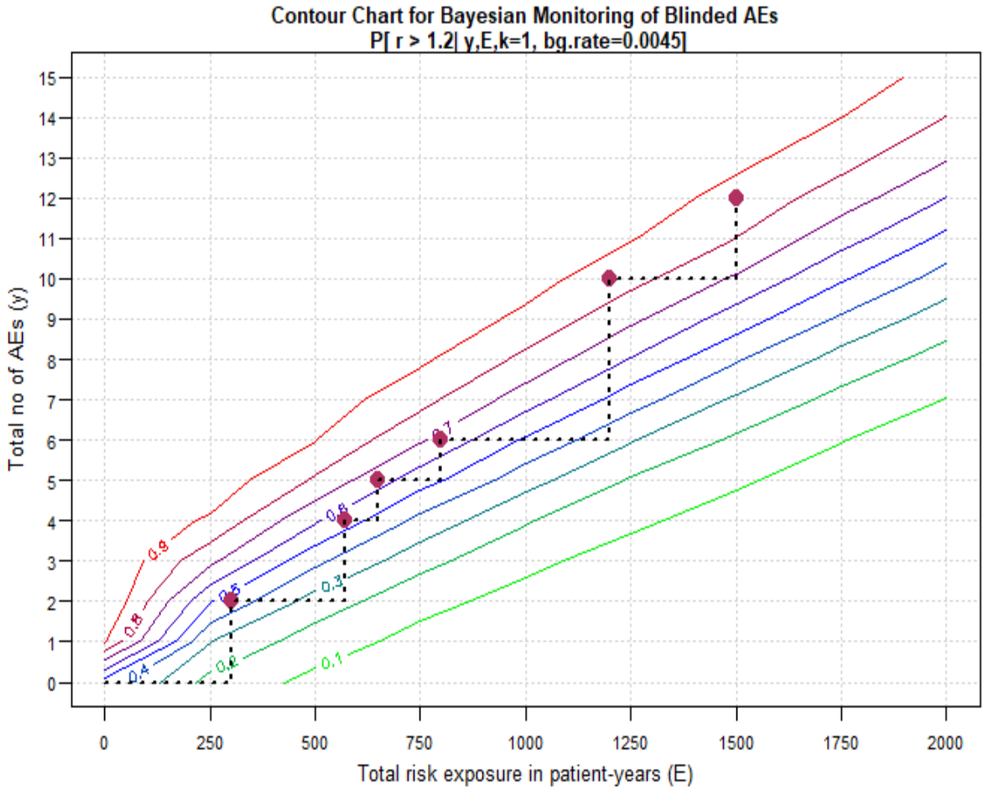
- A planning tool - to prospectively plan to implement
  - The contour plot can be generated beforehand so that the team can use this as a quick planning tool
  - For a given y = events and E = exposure (in pyr), we obtain a contour lines of posterior probability that the relative risk r exceeds a tolerance limit c.
  - Allows to search through various event and exposure combinations
  - To see when we'll make an initial, yet objective, assessment of detecting a signal from blinded safety data

- Can help assess signal without running any analyses
  - Observed event and patient-year (Y, E) can be identified on the contour plot to see the posterior probability of $r > c$



Prospective Planning using Contour Plots

- Contour plot may be further useful to find any trend in the relative risk over time
  - For example, total events of y = 2,4,5,6,10,12 at patient-years 300, 570, 650, 800,1200,1500, respectively, plotted on the contour plot shows an increasing trend of safety signal over time



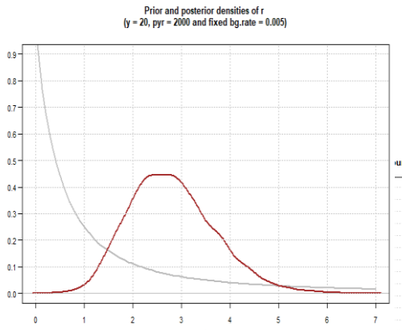
Concluding Remarks

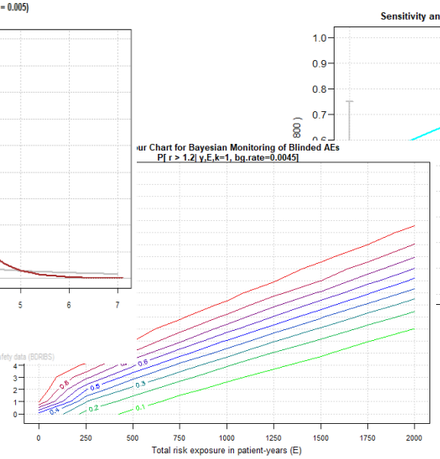
- The proposed approach explicitly models and puts a prior on the signal parameter (relative risk) and thus enables direct inference on signal detection
- Our modeling approach is particularly advantageous as it allows prospective planning and easy implementation of a two-step process
- Some other recent works on blinded monitoring of safety data include Wen et al (2015), Schnell and Ball (2016), Gould and Wang (2017), Ball et al (2020)

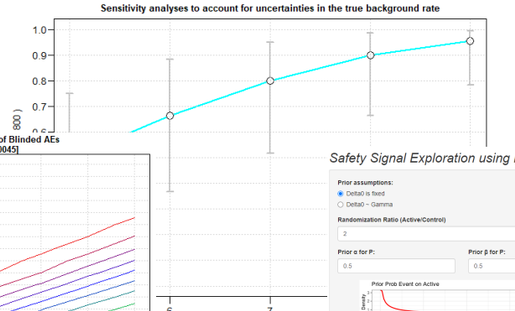
Key References

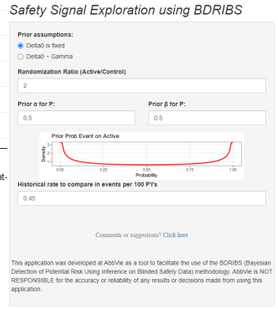
1. Mukhopadhyay, S., Waterhouse, B. and Hartford, A., 2018. Bayesian detection of potential risk using inference on blinded safety data. Pharmaceutical Statistics, 17(6), pp.823-834.
2. Ball, G., Reblin, T., Buchanan, J., Hendrickson, B.A., Lewis, E., Schnell, P.M. and Rockhold, F.W., 2020. A Framework for Safety Evaluation Throughout the Product Development Life-Cycle. Therapeutic innovation & regulatory science, 54(4), pp.821-830.
3. Gould, A.L. and Wang, W.B., 2017. Monitoring potential adverse event rate differences using data from blinded trials: the canary in the coal mine. Statistics in medicine, 36(1), pp.92-104.
4. Schnell, P.M. and Ball, G., 2016. A Bayesian exposure-time method for clinical trial safety monitoring with blinded data. Therapeutic innovation & regulatory science, 50(6), pp.833-838.
5. Wen, S., Ball, G. and Dey, J., 2015. Bayesian monitoring of safety signals in blinded clinical trial data. Annals of Public Health and Research, 2(2), pp.1019-1022.
6. US Department of Health and Human Services Food and Drug Administration. Safety Assessment for IND Safety Reporting. Draft: Guidance for Industry; 2015
7. Goodman, S.N., 1999. Toward evidence-based medical statistics. 2: The Bayes factor. Annals of internal medicine, 130(12), pp.1005-1013.
8. Kass, R.E. and Raftery, A.E., 1995. Bayes factors. JASA, 90(430), pp.773-795.
9. Berger, J.O., 1985. Statistical decision theory and Bayesian analysis. (2nd Ed.) Springer-Verlag.
10. Good, I.J., 1979. Studies in the history of probability and statistics. XXXVII AM Turing's statistical work in World War II. Biometrika, pp.393-396
11. Jeffreys, H., 1961. The theory of probability. Clarendon. Oxford: UK Oxford

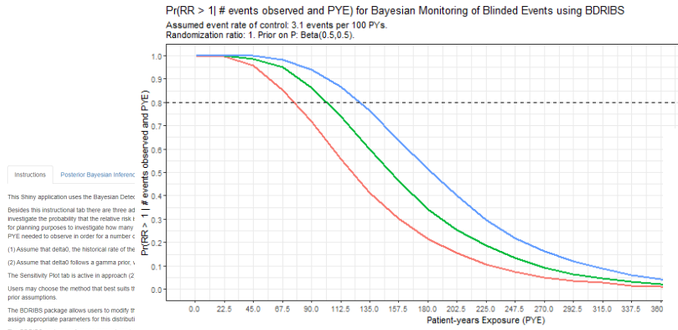
Using R-package and R-shiny dynamic visualization

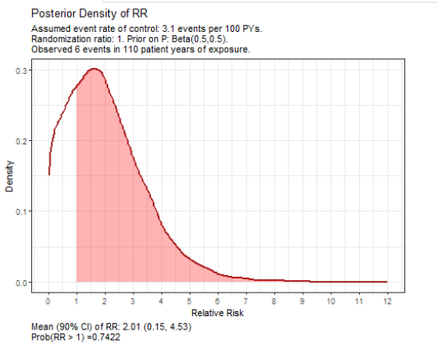












Part II: Implementation

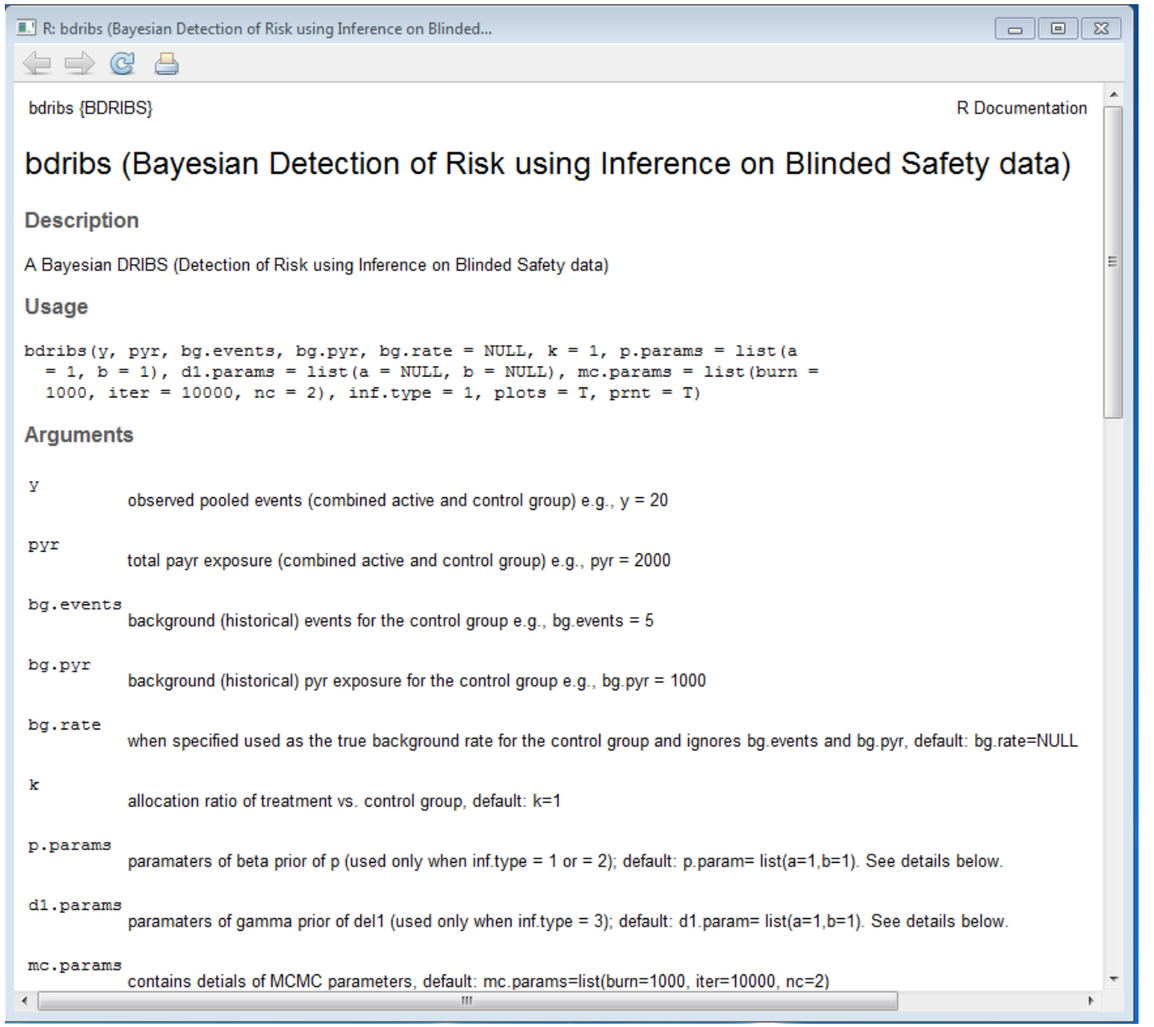
- Bayes computations using BDRIBS R-Package
  - Computing posterior inference
  - Obtaining posterior densities, contour plots, sensitivity plots

- Two-Step Detection Process using BDRIBS R-Shiny application
  - Step1: monitoring and screening for safety signal
  - Step2: verifying signal strength with sensitivity analyses

- Summary, Discussion and Feedback

R-package – BDRIBS

BDRIBS R-package available from CRAN implements relevant Bayesian computations



R-package – BDRIBS: computing posterior inference

- To fit 20 events with pyr=2000 and $p \sim U(0,1)$ with $k = 1$

  - `library(bdribs)`

```
> bd <-bdribs(y=20,pyr=2000, bg.events = 5, bg.pyr = 1000, k=1)
==========================================================================
Printing posterior inference of relative risk (r):
prior is on p ~ beta(a=1, b=1)
y = 20, pyr = 2000, k = 1, bg.rate = 0.005, inf.type = 1
==========================================================================
  mean   5%   95%  P[r>1]  P[r>1.5]  P[r>2]  BF[r>1]  BF[r>1.5]  BF[r>2]
 2.808 1.49 4.378   0.994     0.949   0.818   154.04      27.94     8.96
==========================================================================
```

  - Function ran with default options and produced a default table and plotted estimated posterior density of r.
  - We can do more customized inference using output stored in bd
    - E.g. P[r>1.25] is given by

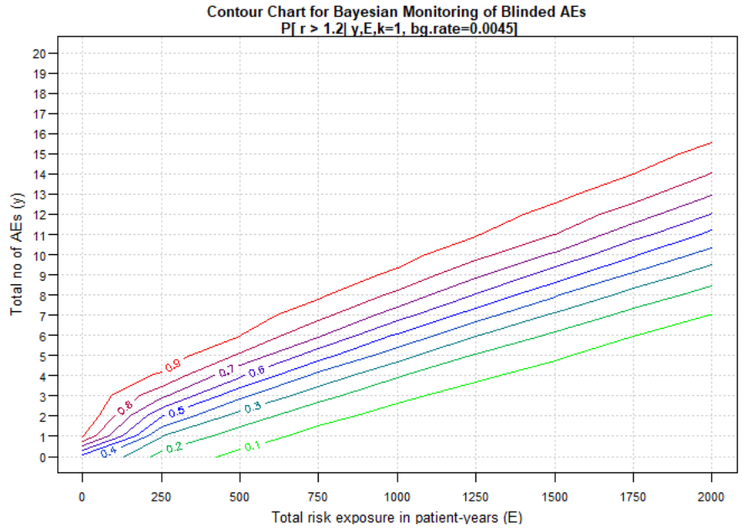
```
> mean(bd$r>1.25)
[1] 0.98045
```

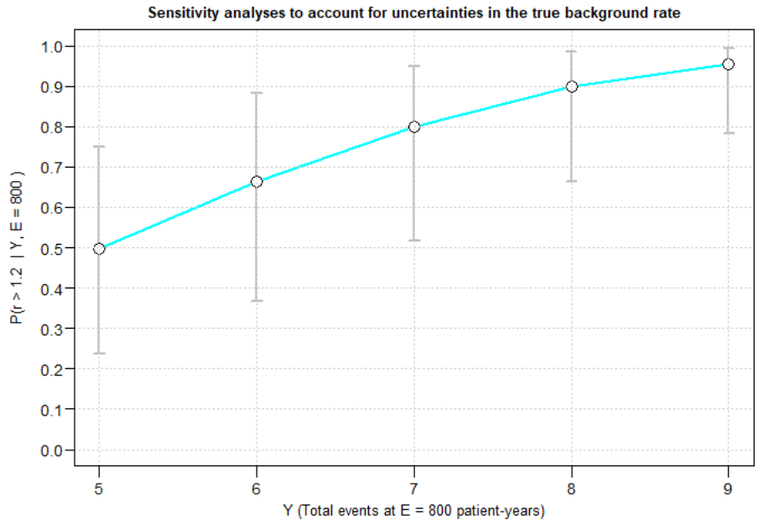
R-package – BDRIBS: contour and sensitivity plots

- Contour and sensitivity plots may be generated desired specifications

```
> bdribs.contour(ymax=20,pyrmax=2000,eincr=250,tol=1.2,k=1, bg.rate=0.0045)
... 5% done
... 11% done
... 16% done
... 21% done
... 26% done
... 32% done
```

```
> bdribs.sensitivity(Y=5:9,pyr=800,k=1, tol=1.2, bg.evnt=18, bg.pyr=4000,bg.ci.coef=0.90)
```



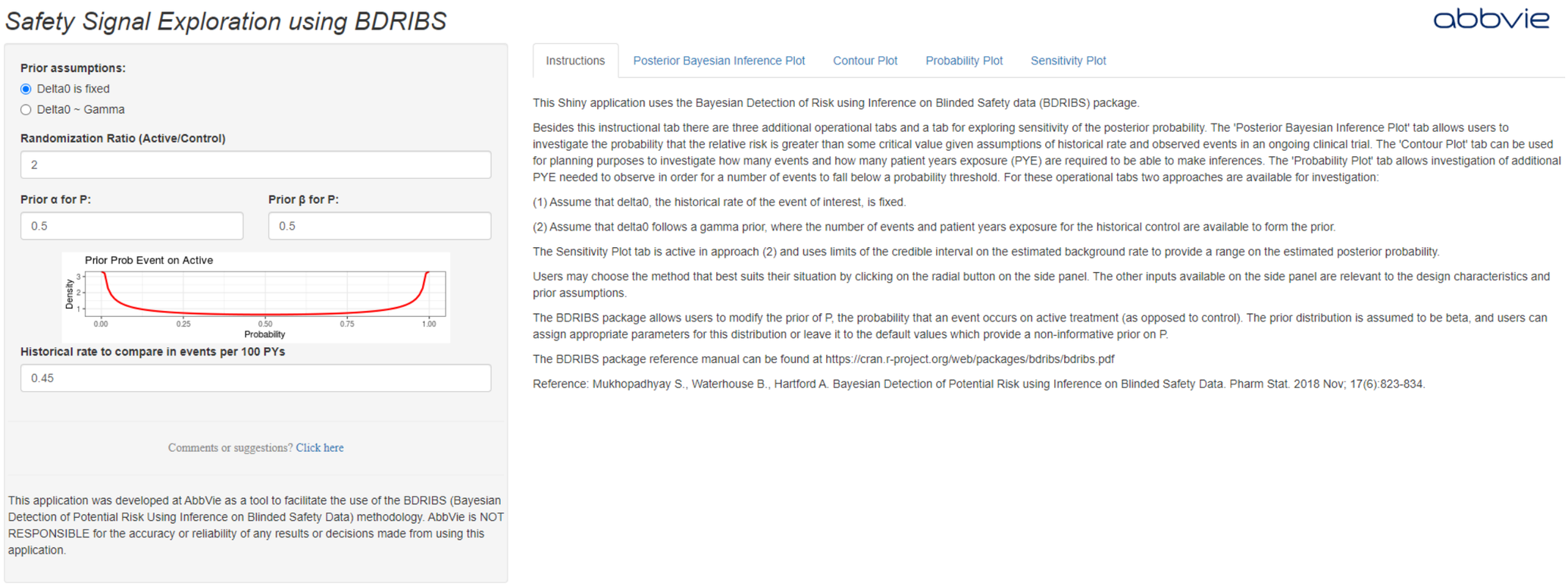


Where BDRIBS fits in the Overall Process of Blinded Safety Signal Detection

Step	Action	Contributor(s)
1	Identification of Anticipated Events and safety topics of interest (STI)	Safety
2	Determination of background reference event rates	Epidemiology
3	Choose appropriate quantitative method and define probability thresholds	Statistics, Safety, Epidemiology
4	Periodic safety data assembly including patient numbers enrolled and patient years of exposure to study drug	Clinical Study Team
5	Identification of event(s) with rates exceeding the pre-designated probability threshold	Programming, Statistics
6	Assessment of totality of evidence for an association of the event with product administration	Statistics, Safety, Clinical, Epidemiology
7	Decision to escalate to safety assessment committee for unblinded review or continue to monitor	Safety Management Team
Multidisciplinary nature of safety signal detection		

BDRIBS R-Shiny App

- To support the use of BDRIBS in planning and to facilitate “live” team discussions an R-Shiny App has been developed
- This App is available at: https://abbviescience.shinyapps.io/BDRIBS/

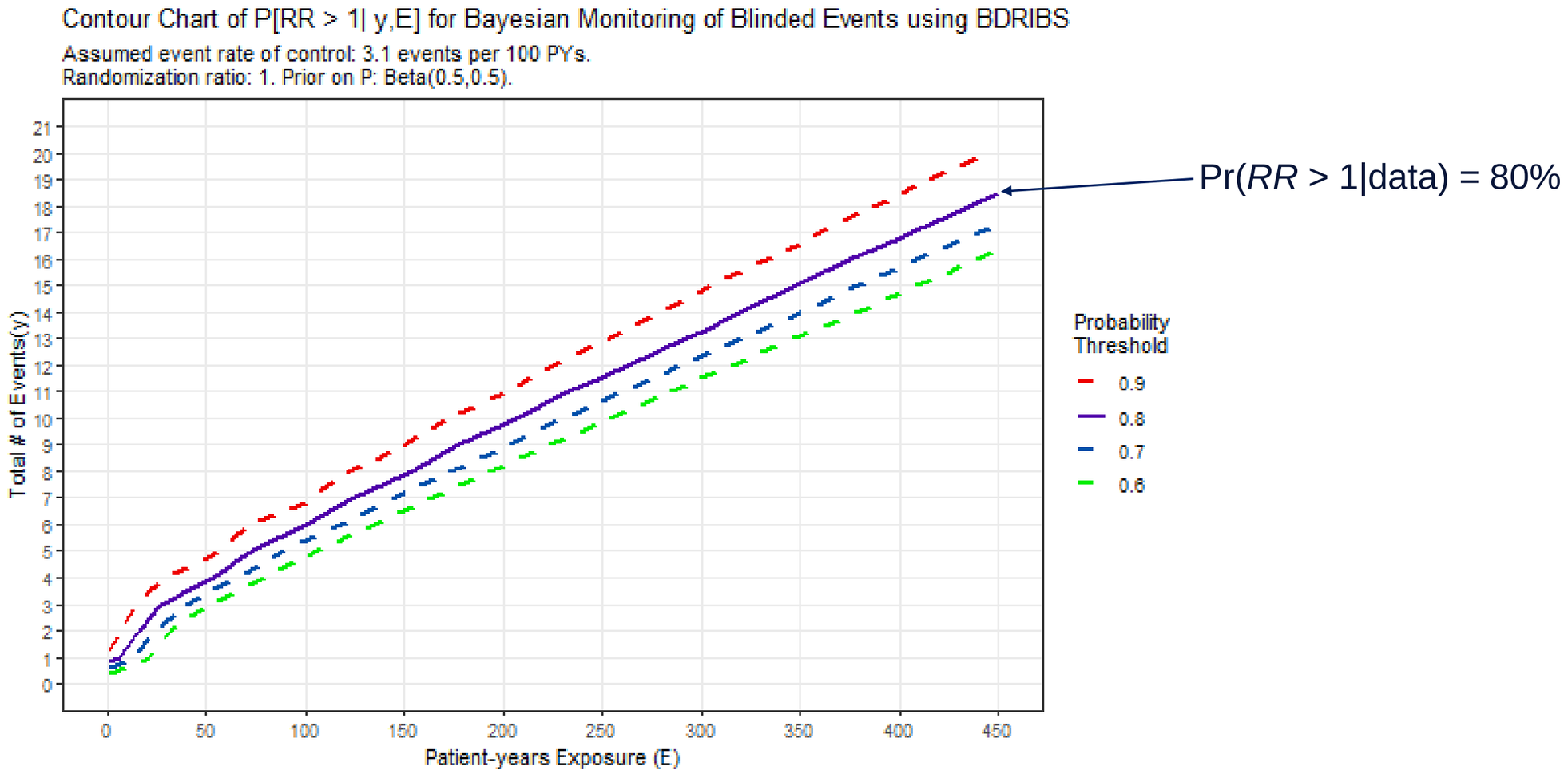


Hypothetical example

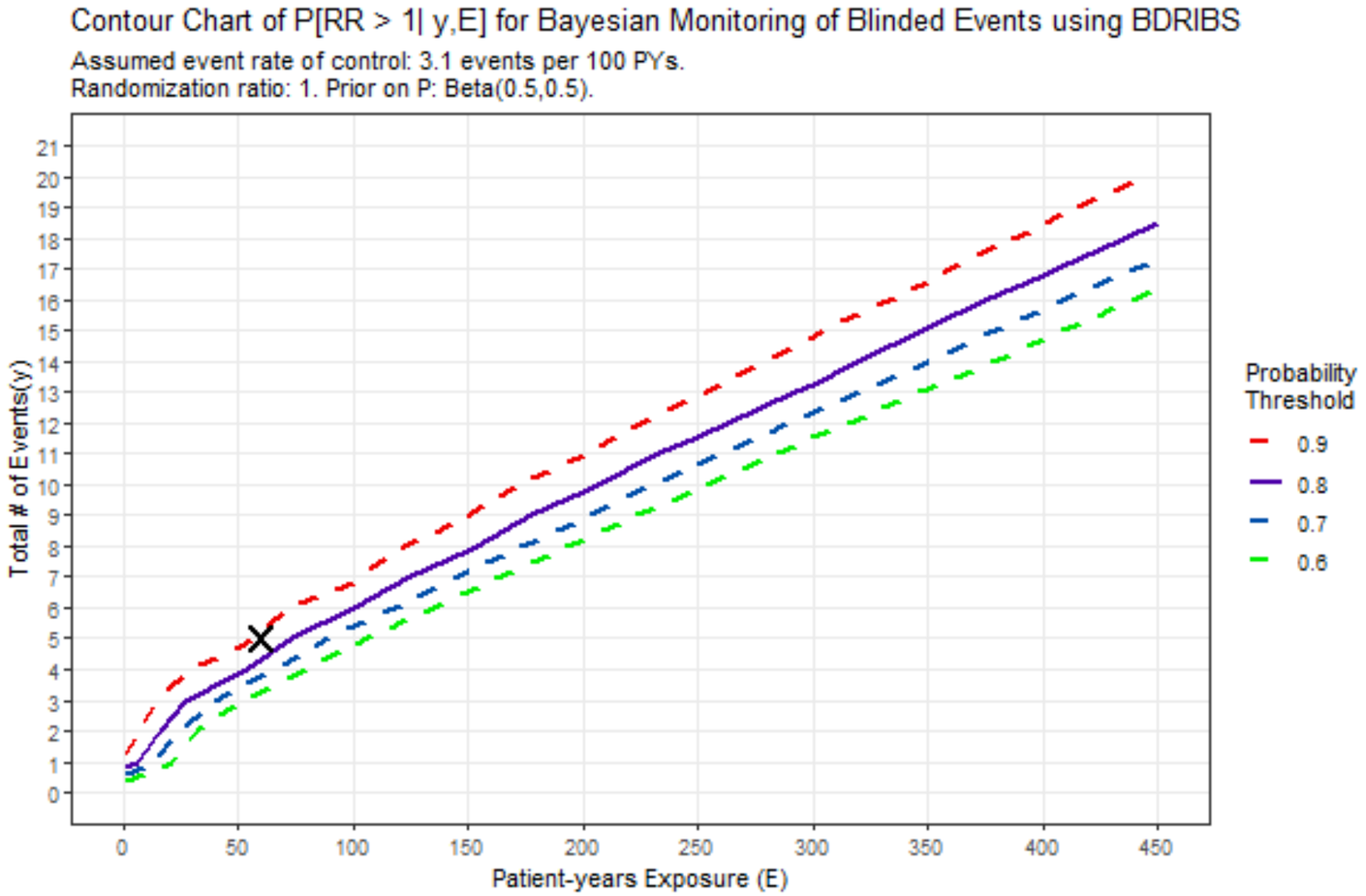
- Indication: Type 2 diabetes
- Double blind, parallel group trial
- Randomization ratio 1:1, active drug or placebo
- Length of treatment: 26 weeks
- N = 900 (450 subjects per arm)
- Cardiovascular safety is also of interest and the team has decided to initiate ongoing blinded safety assessment for MACE.
- The CANVAS trial placebo rate of 3.1 events per 100 PYs for MACE* is assumed for this patient population.
- To date there has been no indication of increased CV risk with the drug being studied, so a relative risk of 1 is of interest and a non-informative prior on P centered on 0.5 is assumed.
- The team decides that if the probability that relative risk exceeds 1 is greater than 80%, further investigation is warranted.

* Neal B, Perkovic V, Mahaffey K W, de Zeeuw D, Fulcher G, Erondu N, Shaw W, Law G, Desai M, Matthews D R. Canagliflozin and Cardiovascular and Renal Events in Type 2 Diabetes. N Engl J Med 2017; 377:644-657

Contour Plot

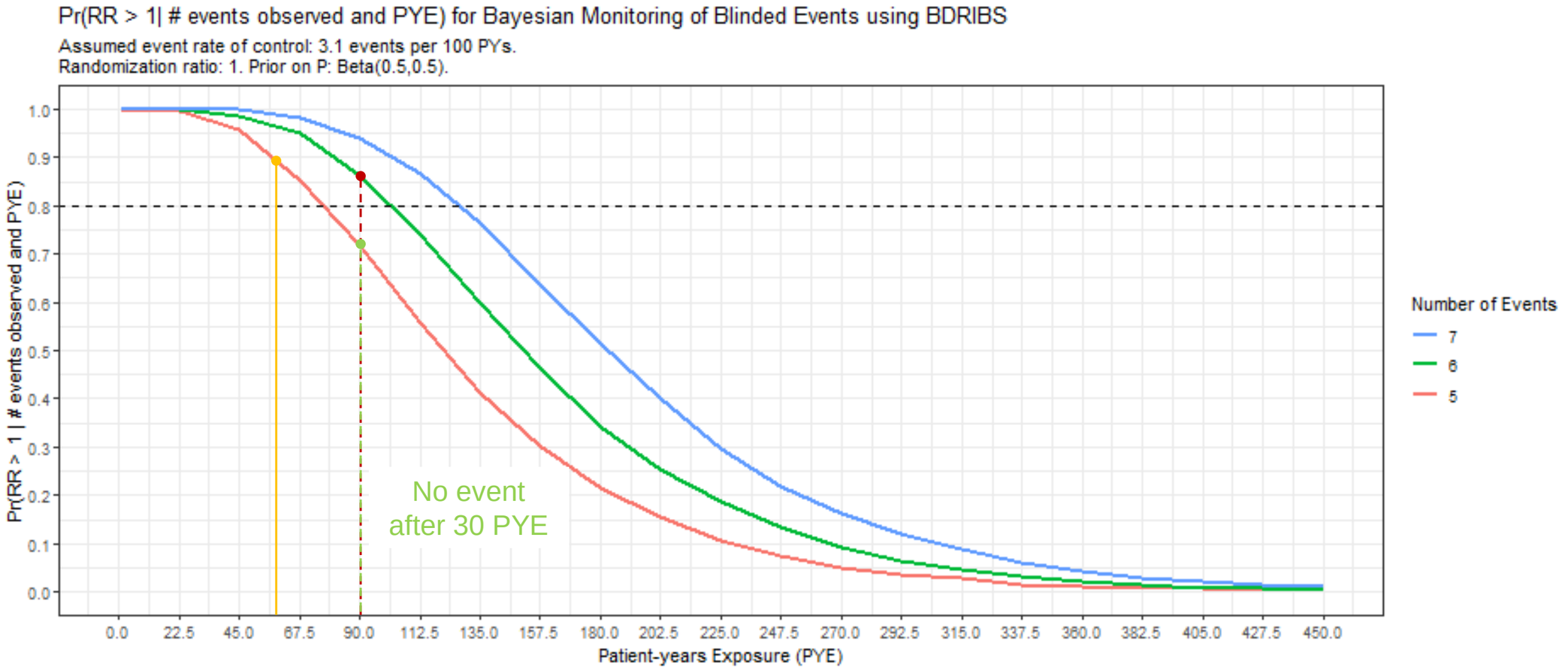


After the trial starts…



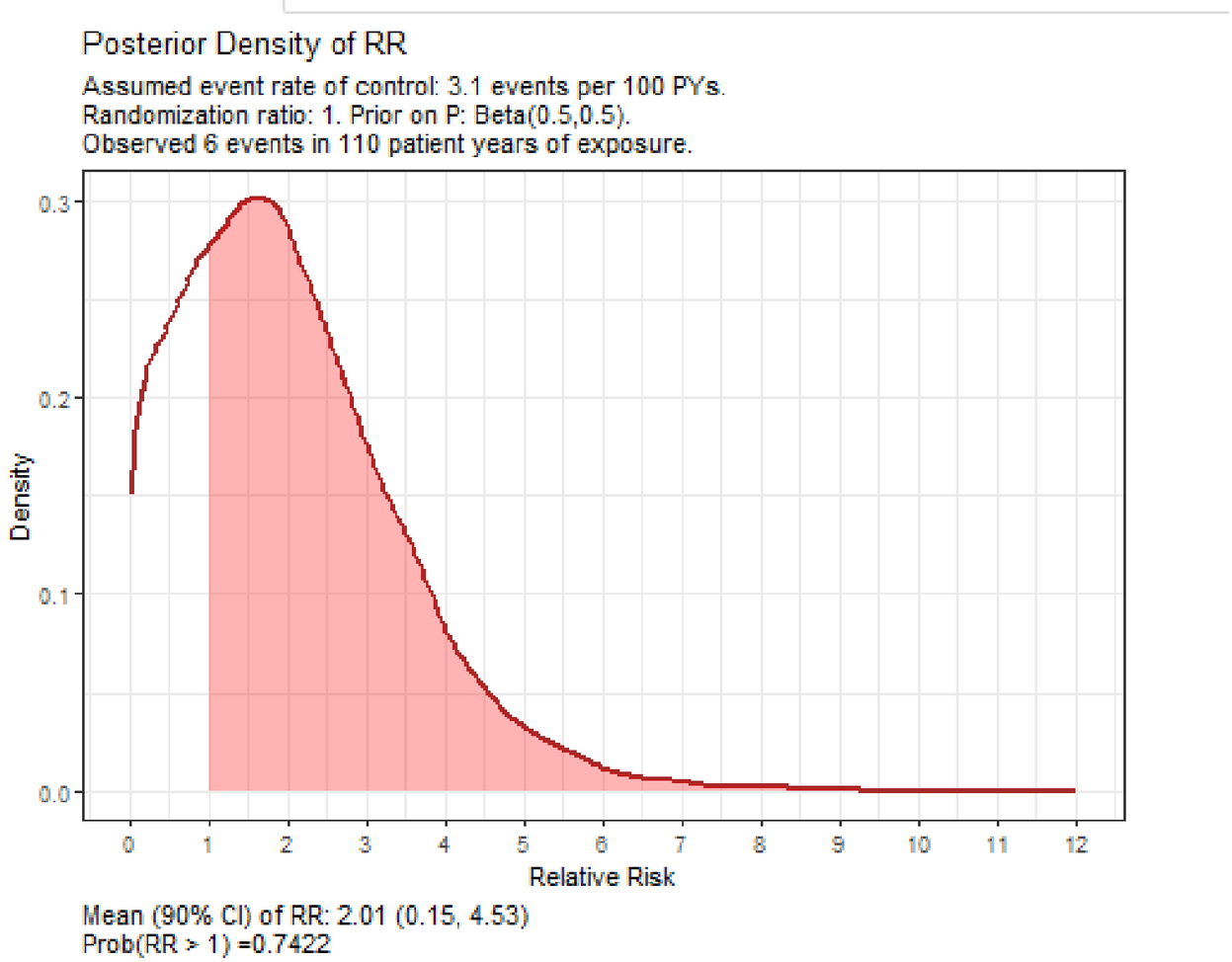
5 events observed after 60 PYs exposure – decide to get the team together to discuss

Probability plots help teams create follow-up rules

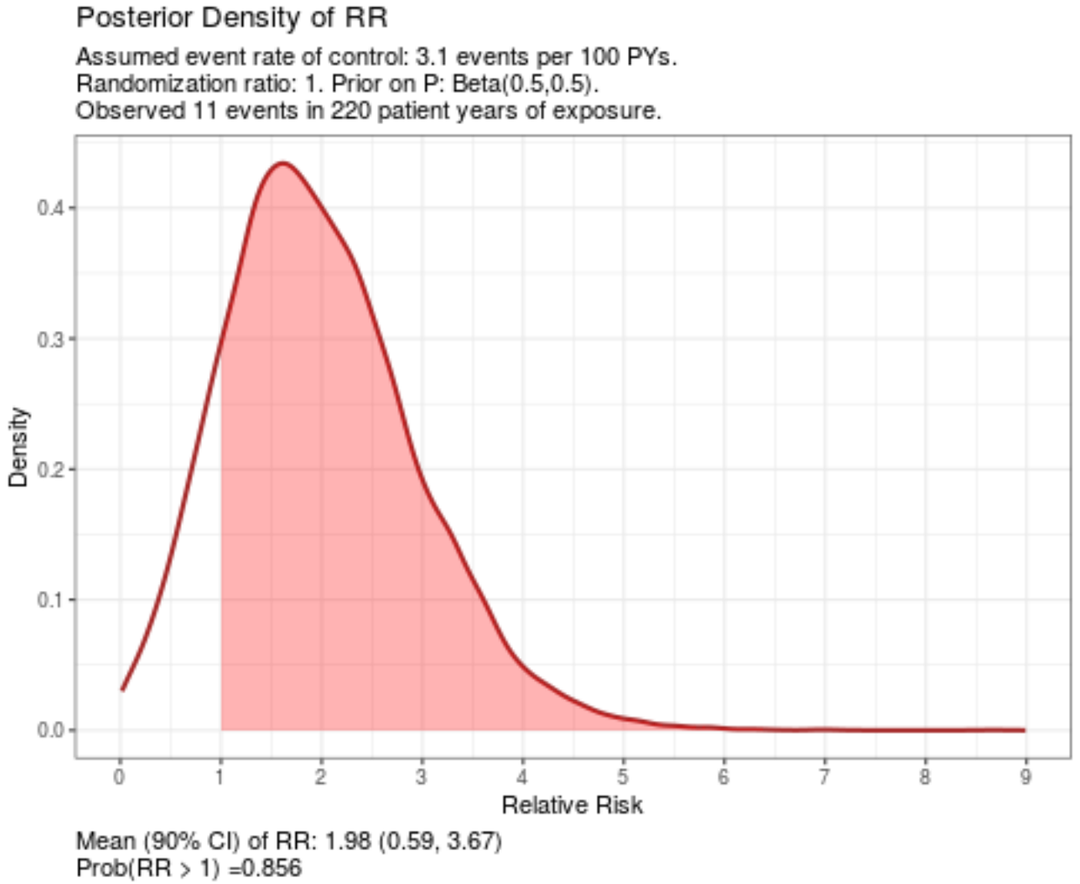


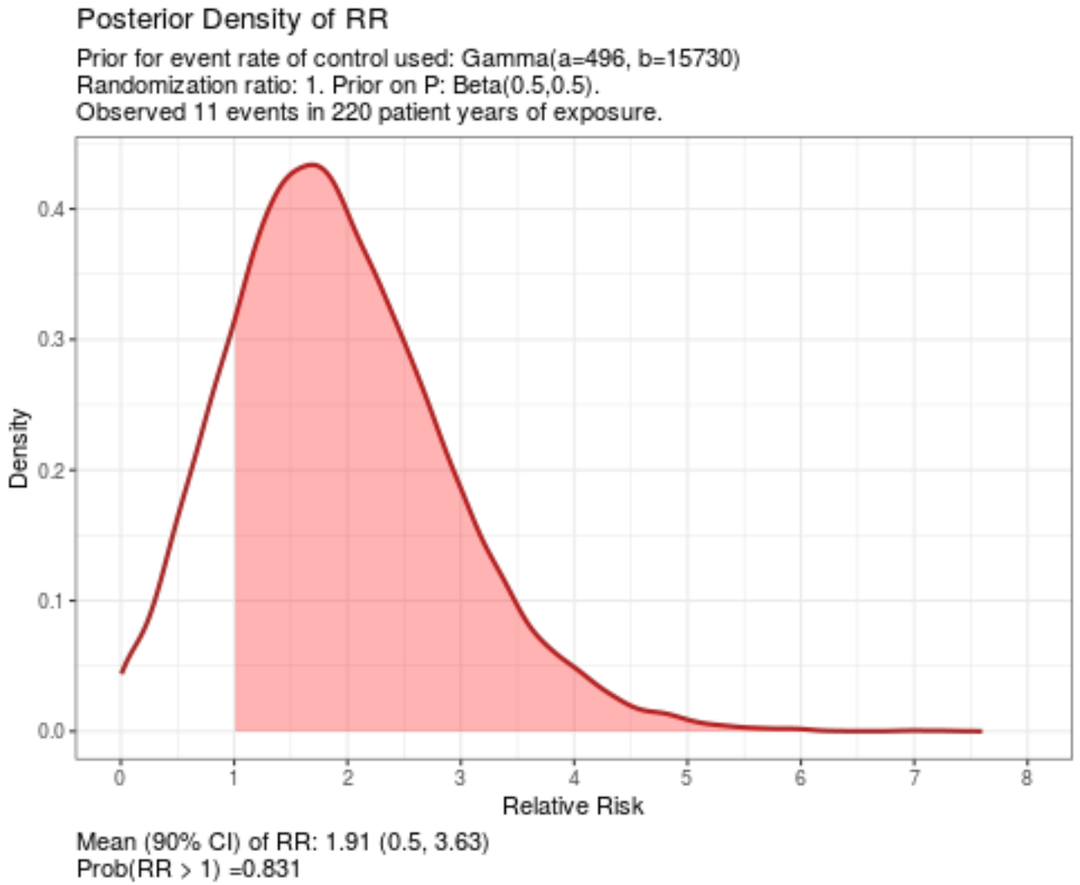
Team decides to wait to see if the next event occurs within the next 30 PYS of exposure

The sixth event occurs after an additional 50 PYs



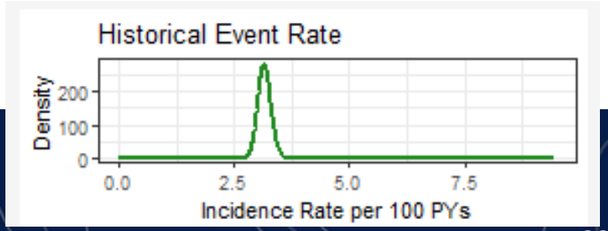
Halfway through the trial 11 events have occurred



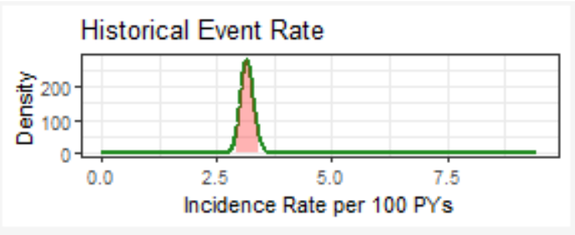


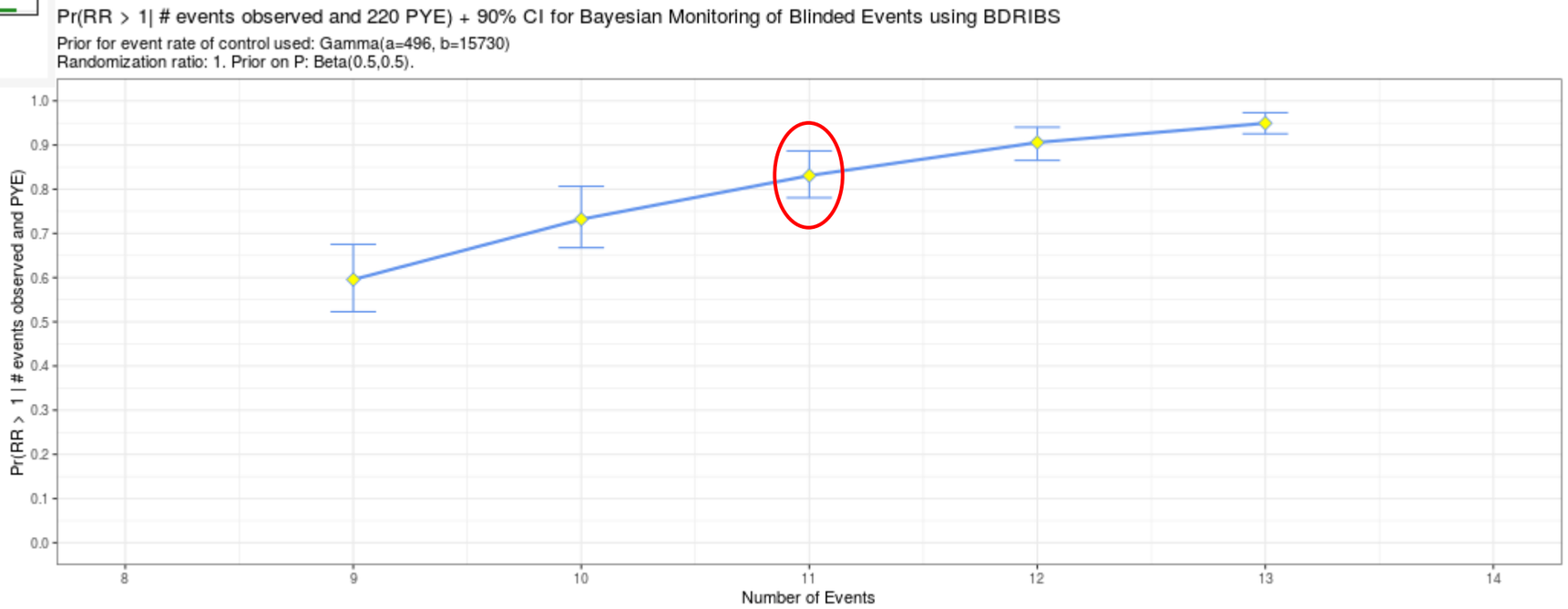
The BDRIBS method looks first at the fixed historical event model and then incorporates the uncertainty of the historical event rate via a gamma prior.

The CANVAS trial event rate of 3.1 E/100 PYs came from 496 events over 15730 PYs.



Investigating the certainty of the probability statement





Considering the 90% CI of the historical event rate, the lower bound of the Pr(RR>1|11 events over 220 PYs) is around the 80% threshold (above 75%). The team decides to refer the event for unblinding.

Summary and Discussion

- BDRIBS (Bayesian Detection of potential Risk using Inference on Blinded Safety data) is a unified Bayesian framework to monitor and detect risk of safety signal from blinded data
  - Explicitly models the signal parameter and thus enables direct inference on signal detection
  - Leverages historical data of the background rates of events of interest
  - Allows easy exploration of the sensitivity of the signal under various assumptions
- Allows prospective planning of the entire implementation and operational procedures
- Provides an objective two-step assessment of new information to detect safety signal
- Easy implementation and operationalization through R-package and dynamic visualization
  - An R package BDRIBS is available on CRAN
  - A BDRIBS R-Shiny application is publicly available from AbbVie Server
- Easily accommodates monitoring combined blinded data from multiple studies
  - The model may be also extended to hybrid settings and/or adjusting for risk factors


abbvie