A Bayesian Sequential Design for COVID-19 Vaccine Trials

DIA BSWG Series

Rajat Mukherjee, PhD

Cytel Inc.

September 18, 2020

Outline

Topics

BCG as a candidate vaccine for reduction of risk of COVID infection

nature > nature reviews immunology > comment > article

Comment | Published: 11 May 2020

BCG-induced trained immunity: can it offer protection against COVID-19?

Luke A. J. O'Neill ✉ & Mihai G. Netea

Nature Reviews Immunology **20**, 335–337(2020) | Cite this article

84k Accesses | **29** Citations | **766** Altmetric | Metrics

Bacillus Calmette–Guérin (BCG) vaccination has been reported to decrease susceptibility to respiratory tract infections, an effect proposed to be mediated by the general long-term boosting of innate immune mechanisms, also termed trained immunity. Here, we discuss the non-specific beneficial effects of BCG against viral infections and whether this vaccine may afford protection to COVID-19.

BCG as a candidate vaccine for reduction of risk of COVID infection

- BCG was developed as a vaccine for TB, however it is known to provide protection against other infectious disease (non-specific effect) such as: respiratory syncytial virus, Yellow fever, herpes simplex virus; human papilloma virus.

BCG as a candidate vaccine for reduction of risk of COVID infection

- BCG was developed as a vaccine for TB, however it is known to provide protection against other infectious disease (non-specific effect) such as: respiratory syncytial virus, Yellow fever, herpes simplex virus; human papilloma virus.
- In clinical studies, BCG vaccination of infants was associated with lower infant mortality, mainly as a result of reduced neonatal sepsis, respiratory infections and fever.

BCG as a candidate vaccine for reduction of risk of COVID infection

- BCG was developed as a vaccine for TB, however it is known to provide protection against other infectious disease (non-specific effect) such as: respiratory syncytial virus, Yellow fever, herpes simplex virus; human papilloma virus.
- In clinical studies, BCG vaccination of infants was associated with lower infant mortality, mainly as a result of reduced neonatal sepsis, respiratory infections and fever.
- A recent placebo controlled study in South Africa showed that BCG revaccination of adults was associated with a 70% reduction in the incidence of acute respiratory infections.

Sponsor's Interest

- Developing a new TB vaccine (similar mechanism of action) but with superior durability of protection - call it nTBV
- Early phase studies have demonstrated the MOA and that nTBV is well tolerated - approval pending on more safety and incidence data over longer follow-ups

Trial Objective

To demonstrate that BCG and/or nTBV vaccine, through innate, trained immune response reduces the incidence of Covid-19 infections during the on-going pandemic.

Trial Objective

To demonstrate that BCG and/or nTBV vaccine, through innate, trained immune response reduces the incidence of Covid-19 infections during the on-going pandemic.

Hypothesis

Superiority of BCG and/or nTBV vaccine in preventing symptomatic Covid-19 infection in exposed health care workers.

Trial Objective

To demonstrate that BCG and/or nTBV vaccine, through innate, trained immune response reduces the incidence of Covid-19 infections during the on-going pandemic.

Hypothesis

Superiority of BCG and/or nTBV vaccine in preventing symptomatic Covid-19 infection in exposed health care workers.

Expected Benefit

- A reduction of at least 45% in incidence rate, powered for 50% reduction

Other Trial Details

- Population: Highly exposed healthcare workers

Other Trial Details

- Population: Highly exposed healthcare workers
- Recruitment rate: Expected 100 per month

Other Trial Details

- Population: Highly exposed healthcare workers
- Recruitment rate: Expected 100 per month
- Randomization: 1:1:1 (BCG:nTBV:Placebo)

Other Trial Details

- Population: Highly exposed healthcare workers
- Recruitment rate: Expected 100 per month
- Randomization: 1:1:1 (BCG:nTBV:Placebo)
- Fixed follow-up of three months

Other Trial Details

- Population: Highly exposed healthcare workers
- Recruitment rate: Expected 100 per month
- Randomization: 1:1:1 (BCG:nTBV:Placebo)
- Fixed follow-up of three months
- Primary Endpoint: Incidence (binary) of Symptomatic COVID-19 illness within 90 days of immunization confirmed by laboratory (RT-PCR)

Other Trial Details

- Population: Highly exposed healthcare workers
- Recruitment rate: Expected 100 per month
- Randomization: 1:1:1 (BCG:nTBV:Placebo)
- Fixed follow-up of three months
- Primary Endpoint: Incidence (binary) of Symptomatic COVID-19 illness within 90 days of immunization confirmed by laboratory (RT-PCR)
- Secondary endpoints include disease severity including the WHO 8-point ordinal scale

Topics

Hypothesis Test

Proportion of subjects getting symptomatic COVID-19 illness within 90 days after immunization:

- BCG arm: π_B; nTBV arm: π_N; Control arm: π_C

Vaccine Efficacy (VE) defined as

$$VE = 1 - \frac{\pi_i}{\pi_C}, \text{ for } i \in (B, N).$$

Hypothesis Test

Proportion of subjects getting symptomatic COVID-19 illness within 90 days after immunization:

- BCG arm: π_B; nTBV arm: π_N; Control arm: π_C

Vaccine Efficacy (VE) defined as

$$VE = 1 - \frac{\pi_i}{\pi_C}, \text{ for } i \in (B, N).$$

Primary Hypothesis:

$$H_0 : \pi_B - \pi_C \geq 0 \quad \cap \quad \pi_N - \pi_C \geq 0$$

VS.

$$H_1 : \pi_B - \pi_C < 0 \quad \cup \quad \pi_N - \pi_C < 0.$$

A Note of Regulatory Requirement

- For vaccine approvals, the point estimate for VE should be at least 0.5 with the lower bound on the alpha-adjusted CI exceeding 0.3 (superiority margin) (Ref: *FDA Guidance, Development and Licensure of Vaccines to Prevent COVID-19, June 2020*)
- This trial is a proof-of-concept trial with limited sample size and follow-up duration - Superiority margin = 0
- Seeking conditional approval based on a successful PoC trial otherwise confirmatory trial to follow if the results are promising.

Expected Sample Size

- Based on a two-sample proportion test with Bonferroni-adjusted significance level of 1.25% for each comparison (BCG vs. Control and nTBV vs. Control). Powering at 90% for an incidence reduction of 50%

Expected Sample Size

- Based on a two-sample proportion test with Bonferroni-adjusted significance level of 1.25% for each comparison (BCG vs. Control and nTBV vs. Control). Powering at 90% for an incidence reduction of 50%
- Assumed control arm proportion over 90 days is 20% (quite high)

Expected Sample Size

- Based on a two-sample proportion test with Bonferroni-adjusted significance level of 1.25% for each comparison (BCG vs. Control and nTBV vs. Control). Powering at 90% for an incidence reduction of 50%
- Assumed control arm proportion over 90 days is 20% (quite high)
- Minimal Sample Size: 945 subjects, with 10% dropouts $N = 1050$ (350 per arm)

Expected Sample Size

- Based on a two-sample proportion test with Bonferroni-adjusted significance level of 1.25% for each comparison (BCG vs. Control and nTBV vs. Control). Powering at 90% for an incidence reduction of 50%
- Assumed control arm proportion over 90 days is 20% (quite high)
- Minimal Sample Size: 945 subjects, with 10% dropouts $N = 1050$ (350 per arm)
- Maximal Sample Size: $N_{max} = 1800$ which will provide 90% power if vaccine efficacy is 0.45 (instead of 0.5) or if the control arm rate if 12% (rather than 20%) but VE is at least 50%.

Expected Sample Size

- Based on a two-sample proportion test with Bonferroni-adjusted significance level of 1.25% for each comparison (BCG vs. Control and nTBV vs. Control). Powering at 90% for an incidence reduction of 50%
- Assumed control arm proportion over 90 days is 20% (quite high)
- Minimal Sample Size: 945 subjects, with 10% dropouts $N = 1050$ (350 per arm)
- Maximal Sample Size: $N_{max} = 1800$ which will provide 90% power if vaccine efficacy is 0.45 (instead of 0.5) or if the control arm rate if 12% (rather than 20%) but VE is at least 50%.

Expected Sample Size

- Based on a two-sample proportion test with Bonferroni-adjusted significance level of 1.25% for each comparison (BCG vs. Control and nTBV vs. Control). Powering at 90% for an incidence reduction of 50%
- Assumed control arm proportion over 90 days is 20% (quite high)
- Minimal Sample Size: 945 subjects, with 10% dropouts $N = 1050$ (350 per arm)
- Maximal Sample Size: $N_{max} = 1800$ which will provide 90% power if vaccine efficacy is 0.45 (instead of 0.5) or if the control arm rate if 12% (rather than 20%) but VE is at least 50%.

Risks Mitigation

- The trial may be under-powered if the control arm incidence (constantly evolving) is lower than assumed
- A Bayesian adaptive design has been proposed

Rationale for a Bayesian Trial Design

- Constantly evolving attack-rate

Rationale for a Bayesian Trial Design

- Constantly evolving attack-rate
- Desire to have a flexible sample size for the PoC trial depending on the control arm rate and VE level ($N_{max} = 1800$ is a guideline), example, if control arm rate is 6% but VE > 55%, the trial could end up with a higher sample size

Rationale for a Bayesian Trial Design

- Constantly evolving attack-rate
- Desire to have a flexible sample size for the PoC trial depending on the control arm rate and VE level ($N_{max} = 1800$ is a guideline), example, if control arm rate is 6% but VE > 55%, the trial could end up with a higher sample size
- Early stopping for futility or dropping one of BCG or nTBV if not efficacious

Rationale for a Bayesian Trial Design

- Constantly evolving attack-rate
- Desire to have a flexible sample size for the PoC trial depending on the control arm rate and VE level ($N_{max} = 1800$ is a guideline), example, if control arm rate is 6% but VE > 55%, the trial could end up with a higher sample size
- Early stopping for futility or dropping one of BCG or nTBV if not efficacious
- Avoid statistical penalty of taking multiple interim looks, using combination test, closed testing procedure, etc. under a frequentist framework

Rationale for a Bayesian Trial Design

- Constantly evolving attack-rate
- Desire to have a flexible sample size for the PoC trial depending on the control arm rate and VE level ($N_{max} = 1800$ is a guideline), example, if control arm rate is 6% but VE > 55%, the trial could end up with a higher sample size
- Early stopping for futility or dropping one of BCG or nTBV if not efficacious
- Avoid statistical penalty of taking multiple interim looks, using combination test, closed testing procedure, etc. under a frequentist framework
- Use totality of evidence available at interim looks to make interim decisions - Use Bayesian predictive power rather than conditional power (Ref: *Dmitrienko-Wang, 2006, SIM*)

Rationale for a Bayesian Trial Design

- Constantly evolving attack-rate
- Desire to have a flexible sample size for the PoC trial depending on the control arm rate and VE level ($N_{max} = 1800$ is a guideline), example, if control arm rate is 6% but VE > 55%, the trial could end up with a higher sample size
- Early stopping for futility or dropping one of BCG or nTBV if not efficacious
- Avoid statistical penalty of taking multiple interim looks, using combination test, closed testing procedure, etc. under a frequentist framework
- Use totality of evidence available at interim looks to make interim decisions - Use Bayesian predictive power rather than conditional power (Ref: *Dmitrienko-Wang, 2006, SIM*)
- Ability to incorporate a new SOC arm and borrow from published external control data

Rationale for a Bayesian Trial Design

- Constantly evolving attack-rate
- Desire to have a flexible sample size for the PoC trial depending on the control arm rate and VE level ($N_{max} = 1800$ is a guideline), example, if control arm rate is 6% but VE > 55%, the trial could end up with a higher sample size
- Early stopping for futility or dropping one of BCG or nTBV if not efficacious
- Avoid statistical penalty of taking multiple interim looks, using combination test, closed testing procedure, etc. under a frequentist framework
- Use totality of evidence available at interim looks to make interim decisions - Use Bayesian predictive power rather than conditional power (Ref: *Dmitrienko-Wang, 2006, SIM*)
- Ability to incorporate a new SOC arm and borrow from published external control data
- Ability to seamlessly extend PoC trial to confirmatory

Final Bayesian Analysis

- Priors: Non-informative $Beta(1,1)$ priors assumed for π_C, π_B and π_M.

Final Bayesian Analysis

- Priors: Non-informative $Beta(1, 1)$ priors assumed for π_C, π_B and π_M.
- Likelihood: Binomial

Final Bayesian Analysis

- Priors: Non-informative $Beta(1,1)$ priors assumed for π_C, π_B and π_M.
- Likelihood: Binomial
- Success criteria at final analysis: The hypothesis test will be carried out only once at the final analysis using posterior distributions for π_i. The success criteria for the i^{th} $(i \in (B, N))$ vaccine arm being

$$Pr\{\pi_i - \pi_C < 0 | \text{ data } \} > \gamma.$$

Final Bayesian Analysis

- Priors: Non-informative $Beta(1,1)$ priors assumed for π_C, π_B and π_M.
- Likelihood: Binomial
- Success criteria at final analysis: The hypothesis test will be carried out only once at the final analysis using posterior distributions for π_i. The success criteria for the i^{th} ($i \in (B, N)$) vaccine arm being

$$Pr\{\pi_i - \pi_C < 0 | \text{ data } \} > \gamma.$$

- Under the Beta-Binomial (conjugate) model the posteriors are also Beta distribution. The posterior probability on differences are calculated using the convolution formula.

Final Bayesian Analysis

- Priors: Non-informative $Beta(1,1)$ priors assumed for π_C, π_B and π_M.
- Likelihood: Binomial
- Success criteria at final analysis: The hypothesis test will be carried out only once at the final analysis using posterior distributions for π_i. The success criteria for the i^{th} ($i \in (B, N)$) vaccine arm being

$$Pr\{\pi_i - \pi_C < 0 | \text{ data } \} > \gamma.$$

- Under the Beta-Binomial (conjugate) model the posteriors are also Beta distribution. The posterior probability on differences are calculated using the convolution formula.
- The success threshold is set to $\gamma = 1 - 0.025/2 = 0.9875$.

Bayesian Adaptive Design

- Multiple interim analyses have been planned with the first interim at around 50% information fraction ($\approx$ 500 completed) with each subsequent interim analysis with 100 additional subjects

Bayesian Adaptive Design

- Multiple interim analyses have been planned with the first interim at around 50% information fraction ($\approx$ 500 completed) with each subsequent interim analysis with 100 additional subjects
- At each interim analysis, Bayesian predictive power (Ref: *Spiegelhalter et al., 2004, Wiley*) with the current cohort will be computed based on the predictive distribution of the binary response conditional on the interim data. This will be done separately for the two vaccine arms

Guidelines for Interim Decision based on PP

Based on the calculated PP the following mutually exclusive interim decisions will be made by the iDMC:

- Stop trial for safety issues; or

Guidelines for Interim Decision based on PP

Based on the calculated PP the following mutually exclusive interim decisions will be made by the iDMC:

- Stop trial for safety issues; or
- Stop trial for futility or drop an arm if PP with planned N (350 per arm) is consistently low

Guidelines for Interim Decision based on PP

Based on the calculated PP the following mutually exclusive interim decisions will be made by the iDMC:

- Stop trial for safety issues; or
- Stop trial for futility or drop an arm if PP with planned N (350 per arm) is consistently low
- If current cohort PP is moderate then either continue to the next interim (if increasing trend) or go to planned end (350 per arm)

Guidelines for Interim Decision based on PP

Based on the calculated PP the following mutually exclusive interim decisions will be made by the iDMC:

- Stop trial for safety issues; or
- Stop trial for futility or drop an arm if PP with planned N (350 per arm) is consistently low
- If current cohort PP is moderate then either continue to the next interim (if increasing trend) or go to planned end (350 per arm)
- If current cohort PP is promising then go to next interim or carry out final analysis with maximum 600 per arm

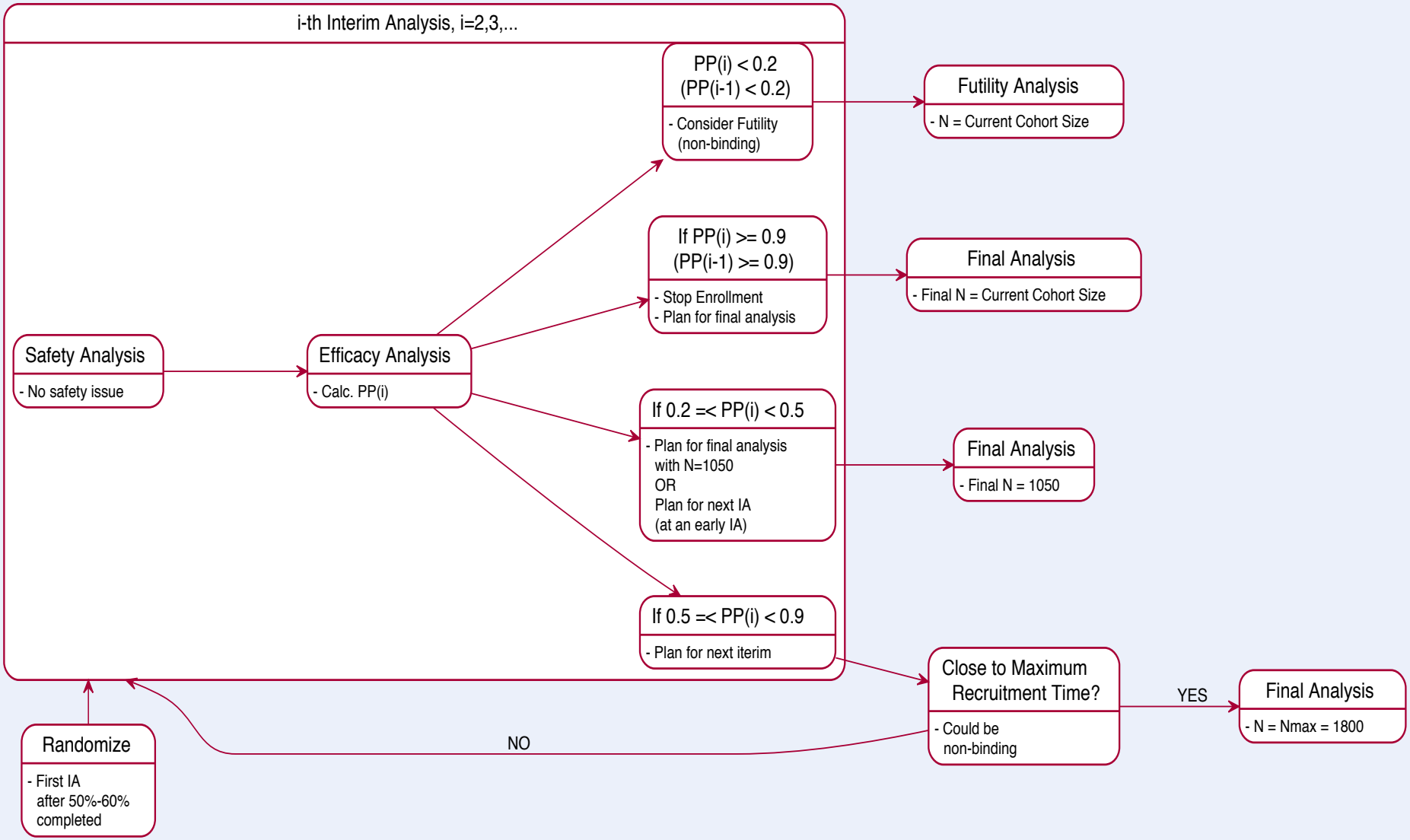
Guidelines for Interim Decision based on PP

Based on the calculated PP the following mutually exclusive interim decisions will be made by the iDMC:

- Stop trial for safety issues; or
- Stop trial for futility or drop an arm if PP with planned N (350 per arm) is consistently low
- If current cohort PP is moderate then either continue to the next interim (if increasing trend) or go to planned end (350 per arm)
- If current cohort PP is promising then go to next interim or carry out final analysis with maximum 600 per arm
- If current cohort PP exceeds the efficacy threshold (say, 0.9) then stop enrollment and carry out final when the last interim cohort has complete follow-up - this is non-binding - may choose to confirm with at least another interim look.

Interim Analysis Plan

Example (used for simulations)



Topics

Predictive Power calculation

Defined as the predictive probability of a successful outcome at the final analysis based on currently enrolled cohort

Current cohort PP

(x, x') is the data of size n from the current cohort, with x the complete data and x' the partial data.

$$PP(n) = \mathbb{E}_{p(x'|x)}\left[\mathcal{I}\left(Pr\{\pi_i - \pi_C < 0 \mid (x, x')\} > \gamma\right)\right],$$

where, the expectation is taken over the posterior predictive distribution of x' given x:

$$p(x'|x) = \int p(\pi|x)L(x', \pi)d\pi.$$

- Actual calculations depend on the nature of interim data (x, x').

PP calc. with binary data only (x' missing)

Algorithm

- Take x as the number of cases (out of m), x' the future number of cases (out of $n - m$)

PP calc. with binary data only (x' missing)

Algorithm

- Take x as the number of cases (out of m), x' the future number of cases (out of $n - m$)
- Beta-Binomial Predictive: $BBP(x'|x, m, n) = \binom{n-m}{x'} \frac{Beta(x+x'+1, n-x-x'+1)}{Beta(x+1, m-x+1)}$

PP calc. with binary data only (x' missing)

Algorithm

- Take x as the number of cases (out of m), x' the future number of cases (out of $n - m$)
- Beta-Binomial Predictive: $BBP(x'|x, m, n) = \binom{n-m}{x'} \frac{Beta(x+x'+1, n-x-x'+1)}{Beta(x+1, m-x+1)}$
- Repeated (10,000) sampling x' from $BBP(x'|x, m, n)$

PP calc. with binary data only (x' missing)

Algorithm

- Take x as the number of cases (out of m), x' the future number of cases (out of $n - m$)
- Beta-Binomial Predictive: $BBP(x'|x, m, n) = \binom{n-m}{x'} \frac{Beta(x+x'+1, n-x-x'+1)}{Beta(x+1, m-x+1)}$
- Repeated (10,000) sampling x' from $BBP(x'|x, m, n)$
- For each sampled x', use (x, x') to get the indicator of success

PP calc. with binary data only (x' missing)

Algorithm

- Take x as the number of cases (out of m), x' the future number of cases (out of $n - m$)
- Beta-Binomial Predictive: $BBP(x'|x, m, n) = \binom{n-m}{x'} \frac{Beta(x+x'+1, n-x-x'+1)}{Beta(x+1, m-x+1)}$
- Repeated (10,000) sampling x' from $BBP(x'|x, m, n)$
- For each sampled x', use (x, x') to get the indicator of success
- Average indicator of success over all 10,000 draws.

PP calc. with binary data only (x' missing)

Algorithm

- Take x as the number of cases (out of m), x' the future number of cases (out of $n - m$)
- Beta-Binomial Predictive: $BBP(x'|x, m, n) = \binom{n-m}{x'} \frac{Beta(x+x'+1, n-x-x'+1)}{Beta(x+1, m-x+1)}$
- Repeated (10,000) sampling x' from $BBP(x'|x, m, n)$
- For each sampled x', use (x, x') to get the indicator of success
- Average indicator of success over all 10,000 draws.

PP calc. with binary data only (x' missing)

Algorithm

- Take x as the number of cases (out of m), x' the future number of cases (out of $n - m$)
- Beta-Binomial Predictive: $BBP(x'|x, m, n) = \binom{n-m}{x'} \frac{Beta(x+x'+1, n-x-x'+1)}{Beta(x+1, m-x+1)}$
- Repeated (10,000) sampling x' from $BBP(x'|x, m, n)$
- For each sampled x', use (x, x') to get the indicator of success
- Average indicator of success over all 10,000 draws.

- Pros: - Speed, used in simulations
- Cons: - Inefficient use of available data, x' could include partial follow-up times but this is extremely computation heavy for simulations. Could be used for actual interim analysis

PP calc. with partial follow-up info

- Use posterior predictive distribution conditional on partial follow-up times
- Use previous algorithm but to draw x_i' use $BBP(x_i' | x(t_i'), m(t_i'), n(t_i'))$, where, t_i' is the partial follow-up time for x_i',
  $x(t_i')$ is the number of observed cases after time t_i',
  $m(t_i')$ is the number surviving (no COVID illness) till time t_i' and
  $n(t_i') - m(t_i')$ is the number of partial follow-ups greater than t_i'.
- This can still be used for simulations, used for the null case

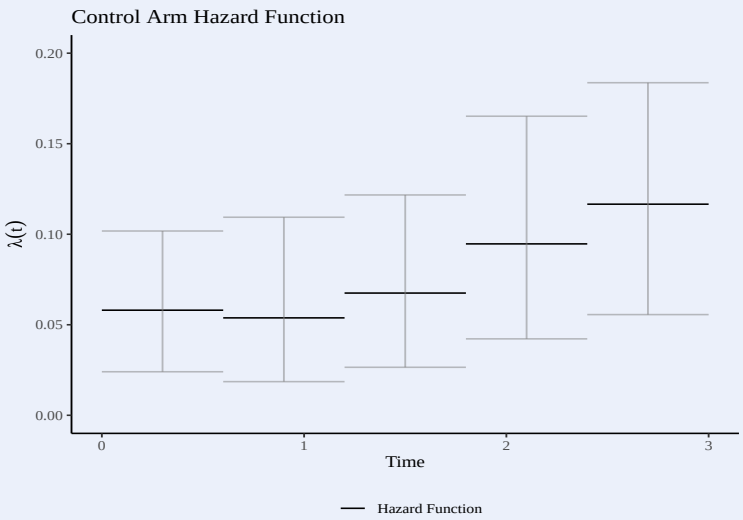
PP calc. with time to event data

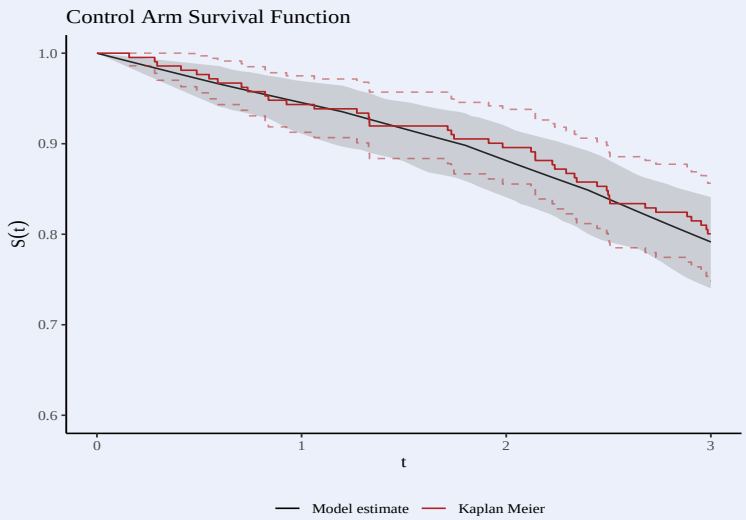
Algorithm (Ref: *Schmidli et al., 2007, Stat. Med.*)

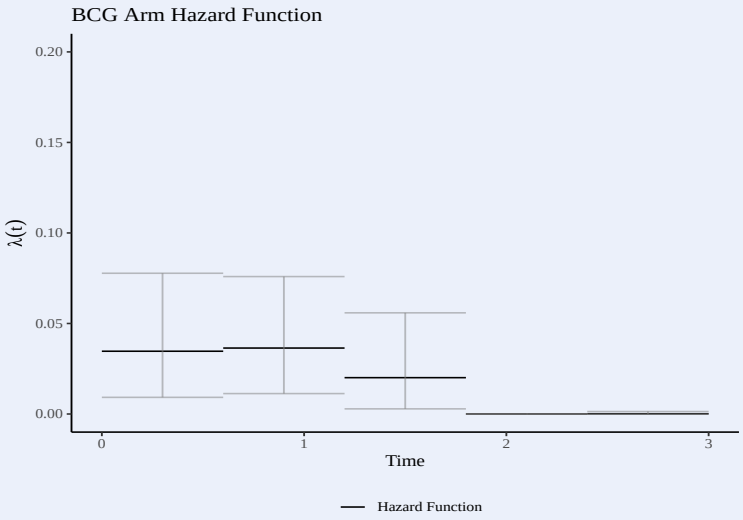
- Model hazard function for each arm using piece-wise exponential distribution.
- Use Gamma-process priors (Ref: *Walker-Barajas, 2002, Sc. J. Stat.*) for the piece-wise constant hazards
- Get the posterior of hazard (or survival) function
- Draw repeatedly (R = 10,000) from this posterior
  - For each draw, predict repeatedly (P = 10,000) infection times (censored at 90 days)
  - Transform predicted data into binary (x') - use (x, x') to compute success indicator
- Average over $R \times P$ draws.

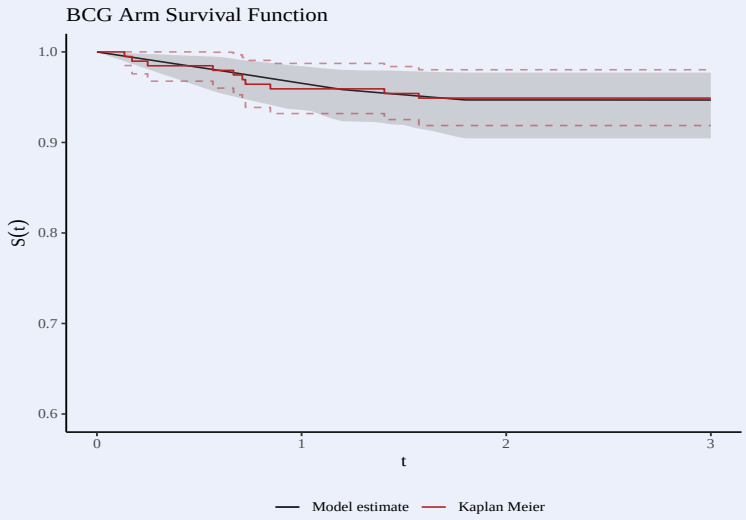
Estimation of Hazard and Survival with Gamma-Process Priors

First Interim Data at 6 months (simulated with mean control arm hazard rate = 0.0637 and HR=0.5), n = 599, n-m = 111 administratively censored, 97 cases









Topics

Simulation Settings

- Recruitment: 100 per month
- Lost-to-follow generated using exponential with rate 1%
- Time to infections generated using piece-wise exponential with average hazard rate 0.067 per-person-month
- First interim look at around 50% information fraction and then every month
- Stopping enrollment early for futility (PP < 0.2) and efficacy (PP ≥ 0.9) based on at two successive PP calculations
- Used $BBP(x'|x, m, n)$ for scenarios under the alternative and $B(x_i'|x(t_i'), m(t_i'), n(t_i'))$ for the null scenario.
- Success criteria: Posterior probability of reduction in incidence proportion compared to control arm > 0.9875.

Operating Characteristics

Simulation Results with Non-binding Futility, at least 2 IAs

Scenario	No. Sims	No. of Interim Mean (SD)	Sample Size Min. Mean (SD) Max.	Duration Min. Mean (SD) Max.	Power/Type-I
null	50000	2.14 (0.68)	622 1043.54 (49.88) 1800	9.22 13.44 (0.5) 21	0.0276
alt.5.5	10000	2.52 (1.14)	619 806.03 (154.78) 1800	9.19 11.06 (1.55) 21	0.9304
alt.5.1	10000	2.63 (1.25)	629 865.18 (170.68) 1800	9.29 11.65 (1.71) 21	0.8444
alt.55.55	5000	2.67 (1.34)	629 847.22 (173.91) 1708	9.29 11.47 (1.74) 20.08	0.8702

- The success threshold can be slightly increased to get type-I error < 0.025.

Topics

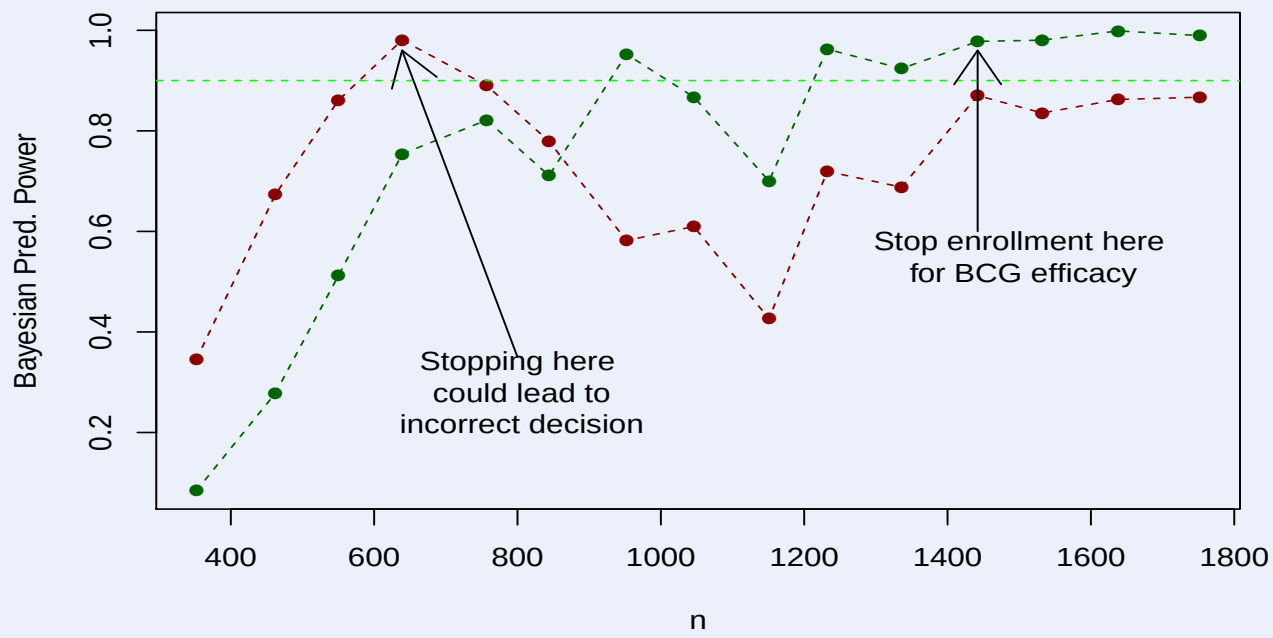
Extension to a Confirmatory Trial

- If the results of this PoC study are promising then the study can be seamlessly (operationally) extended to a confirmatory study with longer follow-up ($\geq$ 24 months) and powered with a superiority margin of 0.3 for VE.
- Inferential seamless option can be also be considered by using the PoC data to construct informative priors.
- However, attack-rates may be different in 10-12 months time. Also the follow-up time will be longer. This will lead to prior-data conflicts
- Options to resolve prior-data conflict must be considered
  - Robust Informative Priors: Adding a weakly-informative mixture component to the informative prior (Ref: *Schmidli et al., 2014, Biometrics*)
  - Use of power prior methodology (Ref: *Ibrahim et al., 2015, SIM*) along with commensurate priors (Ref: *Hobbs et al., 2012, Bayesian Anal.*)

Interim Monitoring

- PP can be calculated at any time during the course of the trial, however, early stopping decision (for efficacy or futility) should be avoided based on early looks (< 50% info. frac.)
- Under the proposed design, it is possible to carry out stopping decisions based on PPs calculated at multiple looks

Fluctuating PPs, simulated data with HR.BCG = 0.5 and HR.nTBV = 0.6



Summary

- Discussed a flexible sequential and adaptive design ideal for rapid establishment of proof-of-concept under uncertainties and logistical constraints
- The design is aimed at facilitating flexible and robust interim decisions
- The predictive power approach can also be used for interim decisions only while the final analysis is carried out using a frequentist test (hybrid design)
- We used non-informative priors here, however, informative priors could be incorporated if historical data is available
- Careful planning and appropriate use of firewalls will reduce operational burden and the chance for operational bias

Thank You!

Email: rajat.mukherjee@cytel.com