Dynamic borrowing of historical data:

Performance and comparison of existing methods based on a case study

D. Dejardin[1], P. Delmar[1], K. Patel[1], C. Warne[1], J. van Rosmalen[2], E. Lesaffre[3].

[1]: F. Hoffmann-La Roche, Basel; [2]: Erasmus MC, Rotterdam; [3]: I-Biostat, KULeuven, Leuven

Outline

Introduction: Antibiotic development

- Number of antibiotics under development low
  - Traditionally large programs required for approval
  - Lack of return on investment
- Growing concerns on antibiotic resistance
  - Unmet medical need (high mortality)
  - Small target population
- Evolving regulatory context
  - Pre-clinical evidence accepted
  - Use of historical data mentioned (FDA guidance for industry 2013)
  - Bayesian methods accepted for devices (FDA guidance for industry 2010)

Introduction: Case study

- Design of phase III comparative study of new agent against pseudomonas aeruginosa (p.a.)
- Target population: Ventilator associated and hospital acquired pneumonia
- Rare condition:
  - 5-10 VAH/HAP per 1,000 hospital admissions
  - 20% caused by p.a.
- Maximum enrollment: 300 subjects total
- Endpoint: 14 days mortality rate (binomial)
- Non-inferiority combo design (maximize safety database)

Motivation for use of historical control

- Subjects are rare
- Need to maximize safety database
  $\Rightarrow$ unbalanced randomization
- Subjects infected, diagnosed, treated, followed up in hospital
  $\Rightarrow$ hope for detailed medical record for historical data

Historical Control

Pocock criteria

1. Treatment in historical control same as randomized control
   $\Rightarrow$ **Met**
2. Historical control form control CT is recent and identical inclusion criteria
   $\Rightarrow$ **Not met**
3. Evaluation of endpoint is the same
   $\Rightarrow$ **Met**
4. Distribution of important subject characteristics are the same
   $\Rightarrow$ **Met**
5. Historical control must be treated in same institution and same investigators
   $\Rightarrow$ **Met**
6. No other indications to expect different outcome
   $\Rightarrow$ **Met**

Historical Control

Pocock criteria

1. Treatment in historical control same as randomized control
   $\Rightarrow$ **Met**
2. Historical control form control CT is recent and identical inclusion criteria
   $\Rightarrow$ **Not met**
3. Evaluation of endpoint is the same
   $\Rightarrow$ **Met**
4. Distribution of important subject characteristics are the same
   $\Rightarrow$ **Met**
5. Historical control must be treated in same institution and same investigators
   $\Rightarrow$ **Met**
6. No other indications to expect different outcome
   $\Rightarrow$ **Met**

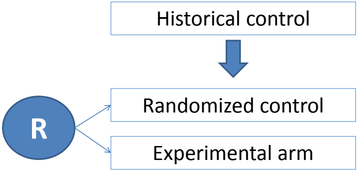
- **Very strict criteria**
- **No guaranty that prior match**

Historical control and randomized control

- Concern: Uncontrolled factors may impact validity of historical control

Historical control and randomized control

- Concern: Uncontrolled factors may impact validity of historical control
- Use Small randomized control to check compatibility of historical control



Dynamic Bayesian Borrowing

Goal of methods

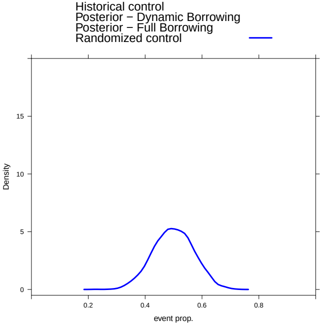
Increase precision when compatible, control bias when not compatible

Dynamic Bayesian Borrowing

Goal of methods

Increase precision when compatible, control bias when not compatible

Compatible historical data

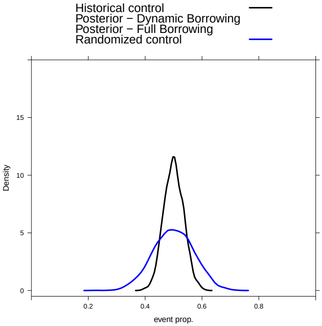


Dynamic Bayesian Borrowing

Goal of methods

Increase precision when compatible, control bias when not compatible

Compatible historical data

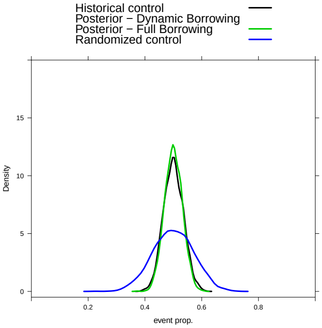


Dynamic Bayesian Borrowing

Goal of methods

Increase precision when compatible, control bias when not compatible

Compatible historical data

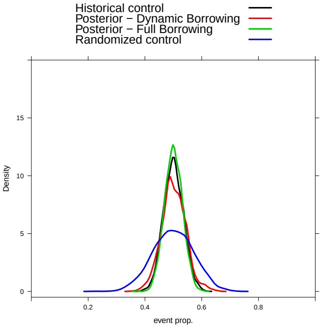


Dynamic Bayesian Borrowing

Goal of methods

Increase precision when compatible, control bias when not compatible

Compatible historical data



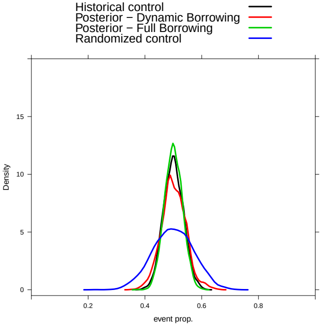
- Increase in precision

Dynamic Bayesian Borrowing

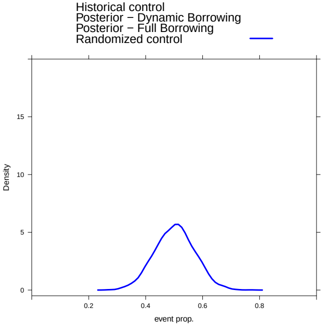
Goal of methods

Increase precision when compatible, control bias when not compatible

Compatible historical data



Incompatible historical data



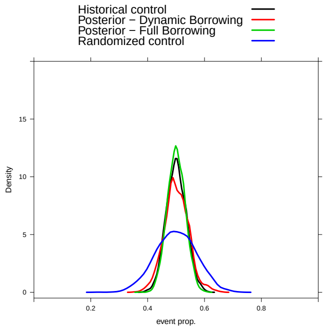
- Increase in precision

Dynamic Bayesian Borrowing

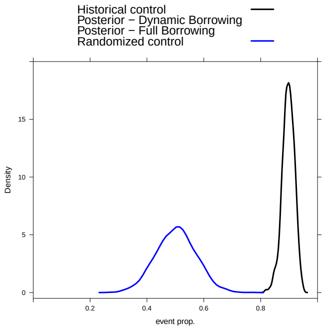
Goal of methods

Increase precision when compatible, control bias when not compatible

Compatible historical data



Incompatible historical data



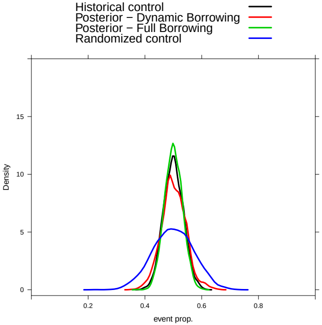
- Increase in precision

Dynamic Bayesian Borrowing

Goal of methods

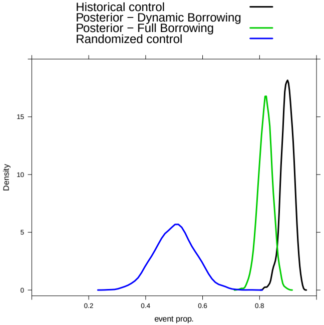
Increase precision when compatible, control bias when not compatible

Compatible historical data



- Increase in precision

Incompatible historical data



Full Borrowing

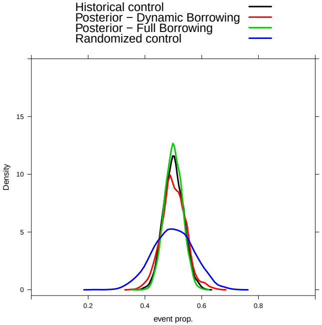
- Increase in Bias

Dynamic Bayesian Borrowing

Goal of methods

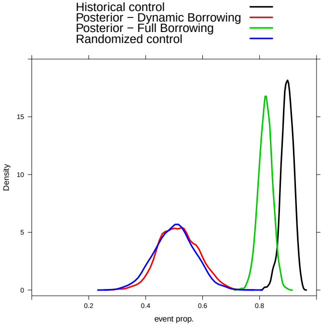
Increase precision when compatible, control bias when not compatible

Compatible historical data



- Increase in precision

Incompatible historical data



Dynamic Borrowing

- Bias controlled

Methods

- **Normalized power prior**
- **Robust mixture prior**
- **Commensurate prior**

Normalized power prior

Prior for event rate p:

$$\pi^P(p,\theta|H) \propto \underbrace{\frac{1}{C(\theta)}}_{\text{Normalizing cst}} \Big[\underbrace{L^H(p)}_{\text{Historical data}} \Big]^{\theta} \underbrace{\pi_v(p)}_{\text{vague prior}} \underbrace{\pi_v(\theta)}_{\text{vague prior for } \theta}$$

- Historical data (H) prior raised to power θ
- $\theta \in [0,1] =$ measure of compatibility
    - $\theta = 0 \Rightarrow$ No Borrowing
    - $\theta = 1 \Rightarrow$ FULL Borrowing
- θ jointly estimated with p

Robust Mixture prior

Prior for event rate p:

$$\pi^{mx}(p|H) = w \underbrace{\pi^{H}(p)}_{\text{Historical data}} + (1-w) \underbrace{\pi^{v}(p)}_{\text{vague prior}}$$

- Weights w are pre-specified
- Determined through simulations
- Weights are updated in posterior
- Random weights possible, but do not depend on data

Commensurate prior

$$\pi^C(p, p_h, \sigma | H) \propto \underbrace{L^H(p_h)}_{\text{Historical data}} \underbrace{\psi(p, p_h, \sigma)}_{\text{link function}} \underbrace{\pi_v(p_h, \sigma)}_{\text{vague prior}}$$

- Separate parameters for randomized (p) and historical (p_h)
- Connected through a link function (distribution)
  - Mean of link distribution $= p_h$
  - Variance $= \sigma =$ measure of compatibility
    - High variance $\Rightarrow$ Low compatibility
    - Low variance $\Rightarrow$ high compatibility

Note on method

All methods depend on parameters : Need for calibration

- Robust Mixture Prior: pre-specified weight w
- Commensurate prior: Prior for variance $\pi_v(\sigma)$
- Normalized power prior: Prior for power parameter $\pi_v(\theta)$
  $\Rightarrow$ Natural choice: Jeffreys' prior Beta(1/2,1/2)

$\Rightarrow$ **All methods calibrated on NPP**
$\Rightarrow$ **On maximum type I error**

Simulations

Goal of simulations

Compare methods:

- Impact of drift on type I error
- Increase in power
- Against
  - Frequentist test: No borrowing, just randomized data
  - Full borrowing: Simple Bayesian analysis (pooled analysis)

Here:

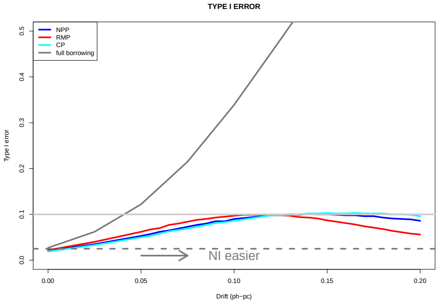
- Design stage: Best assumptions on control and experimental mortality rate
- Control collected at site opening:
  - Unknown historical rate
  - Historical rate to be simulated as well

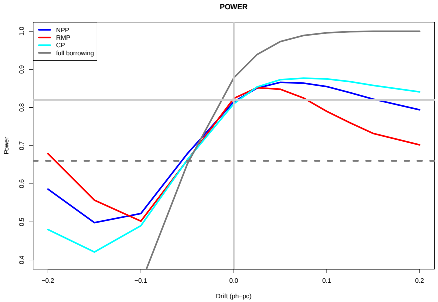
Simulation settings

- Non-inferiority test for mortality rate ($\searrow$ better)
- Rate in randomized control: $p_c = 25\%$
- Non inferiority margin $= 12.5\%$

  Positive test if 95% CI of difference $p_e - p_c < 12.5\%$
- Rate in Incompatible historical control: $p_h = 37.5\%$
- Rate in compatible historical control: $p_h = 25\%$
- Sample size:
  - Historical control : 200
  - Randomized control: 100
  - Experimental arm: 200

Results





- Maximum α set to 10% (following calibration)
- Power gain: 12% (Power = 82%)
- All methods similar

Comparison with frequentist

- Previous plot: All methods calibrated to max(α)$=$ 0.10

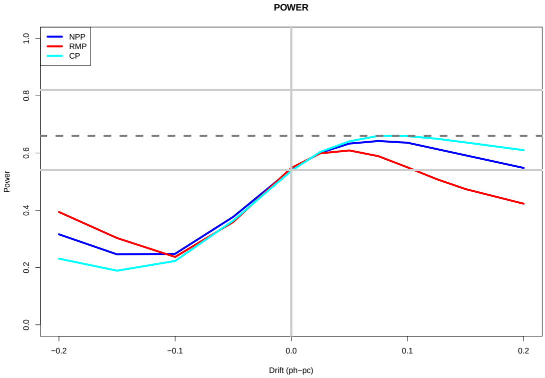
  Except "Frequentist" α=0.025

- 2 options align α

  1. Lower Dynamic borrowing to maximum $\alpha \leq 0.025$
  2. Allow frequentist to have $\alpha = 10\%$

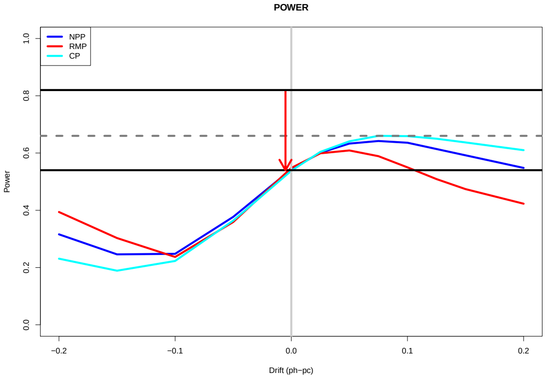
Option 1

- **Dynamic borrowing to maximum** $\alpha \leq 0.025$
- Change width of CI (methods use 95% CI):
  Width: 95% $\Rightarrow$ 99%



Option 1

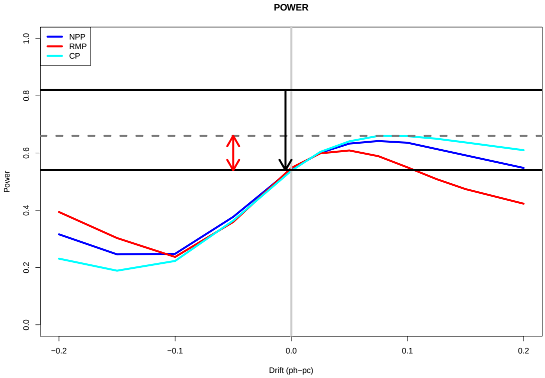
- **Dynamic borrowing to maximum $\alpha \leq 0.025$**
- Change width of CI (methods use 95% CI):
  Width: 95% $\Rightarrow$ 99%



- Power reduced from 0.82 to 0.54

Option 1

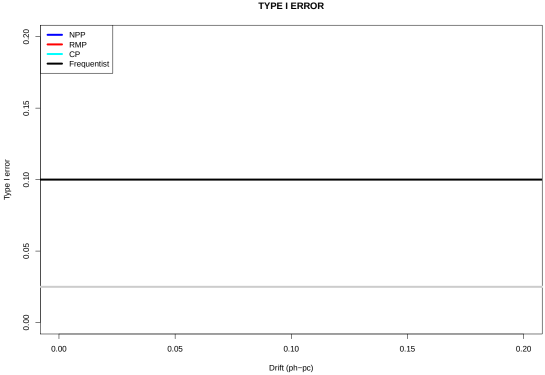
- **Dynamic borrowing to maximum** $\alpha \leq 0.025$
- Change width of CI (methods use 95% CI):
  Width: 95% $\Rightarrow$ 99%



- Power reduced from 0.82 to 0.54
- Power of dynamic borrowing (0.54) lower than no historical data (0.66)

Option 2

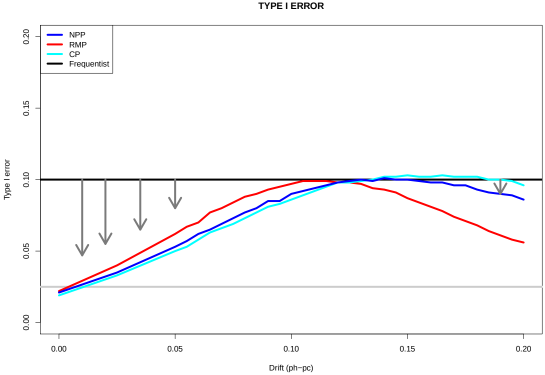
- **Frequentist** $\alpha = 10\%$
- Power of frequentist method increases $\rightarrow 0.87$
- **Frequentist higher power than Dynamic borrowing (0.82)**



- Frequentist α constant $= 0.1$ over drift

Option 2

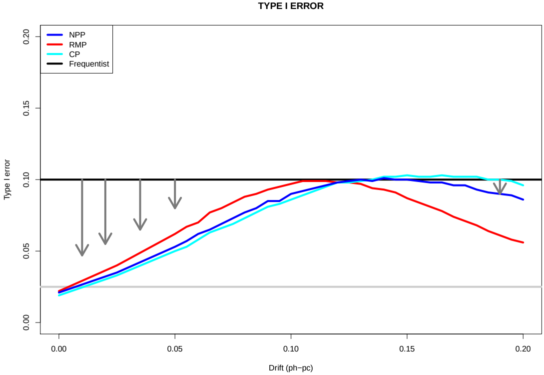
- **Frequentist** $\alpha = 10\%$
- Power of frequentist method increases $\rightarrow$ 0.87
- **Frequentist higher power than Dynamic borrowing (0.82)**



- Frequentist α constant $= 0.1$ over drift
- Dynamic borrowing α mostly below 0.1

Option 2

- **Frequentist** $\alpha = 10\%$
- Power of frequentist method increases $\rightarrow$ 0.87
- **Frequentist higher power than Dynamic borrowing (0.82)**



- Frequentist α constant $= 0.1$ over drift
- Dynamic borrowing α mostly below 0.1

$\Rightarrow$ **Dynamic borrowing methods have lower** α

Discussion

- **Can Dynamic borrowing replace a frequentist analysis (gain power)**
  - **NO,** when strict control of α required
  - **YES** , if some increase is allowed
    - Makes sense: Historical data = trustworthy source of data
    - Type I error inflation depends on historical data
    - Risk (α/power) of historical data is limited but not suppressed!
- **Benefits of Dynamic borrowing**
  - Limit bias, type I error compared to full borrowing in case incompatible data
  - Mixture prior is best (Simple - needs optimization - fast)
  - Commensurate difficult to implement
    Linked to variance parameter controlling for compatibility

References

- Pocock, S. J. *The combination of randomized and historical controls in clinical trials* Journal of Chronic Diseases , **1976**, 29, 175 - 188
- Hobbs, B. P.; Carlin, B. P.; Mandrekar, S. J. & Sargent, D. J. *Hierarchical Commensurate and Power Prior Models for Adaptive Incorporation of Historical Information in Clinical Trials*, Biometrics, **2011**, 67, 1047-1056
- Duan, Y.; Ye, K. & Smith, E. P. *Evaluating water quality using power priors to incorporate historical information*, Environmetrics, John Wiley & Sons, Ltd., **2006**, 17, 95-106
- Neuenschwander, B.; Branson, M.& Spiegelhalter, D. J. *A note on the power prior*, Statistics in Medicine, John Wiley & Sons, Ltd., **2009**, 28, 3562-3566
- Schmidli, H.; Gsteiger, S.; Roychoudhury, S.; O'Hagan, A.; Spiegelhalter, D. & Neuenschwander, B. *Robust meta-analytic-predictive priors in clinical trials with historical control information*, Biometrics, **2014**