

Bayesian Augmented Control:

Borrowing from Historical Control information
to enhance a trial



Kert Viele and Scott Berry
FDA Industry Workstop
September 20th, 2011

Introduction

- In orphan indications one often is faced with minimal resources to conduct a trial.
- Maximizing the information obtained from the trial is therefore vital. Any information from outside the trial should be used efficiently
- Bayesian Augmented Control methods allow for information on the control arm to be supplemented by previous trials, without dominating the results of the current trial.

Example

- Oncology example (time to event) – resources are only available for 100 patients. (side details include 2 pts/wk accrual, 26 wk followup on last patient)
- How to allocate?
 - Single arm trial? Must assume control rate...
 - Two arm trial with 50-50 allocation? Not many patients on the treatment arm...
 - Two arm trial with 25-75 allocation (more on treatment)? Not many on control, need control for comparison...

Suppose you have historical data

- Suppose 4 previous trials have been conducted in this area using the standard of care (control) arm. The trials had data
 - 37 events, 1233 weeks exposure, hazard 0.0300
 - 7 events, 211 weeks exposure, hazard 0.0332
 - 104 events, 3811 weeks exposure, hazard 0.0273
 - 71 events, 2218 weeks exposure, hazard 0.0320
- Trails all have hazards around 0.03, but also show variability.
- Can we use this to inform the control arm?

Comparing Possible Trials

- Will examine control in current trial to be any of hazards=0.025,0.0275,0.030,,0.0325,0.035 with treatment hazard ratios of HR=1,0.9,0.8,0.7,0.6
- Success defined as posterior probability $\Pr(\text{HR} < 0.9)$ greater than 90%.
- Evaluate operating characteristics of type I error and power at the 25 possibilities (crossing control hazard and HR)

Trial 1 – Ignore history, go 50-50

- As a baseline, suppose we ignore the historical data and run a 2 arm, equal allocation design

Type I error 10%

	1.00	0.90	0.80	0.70	0.60
Better	0.1015	0.2071	0.3517	0.5557	0.7692
better	0.1037	0.2063	0.3606	0.5718	0.7853
Same	0.1064	0.2071	0.3786	0.5900	0.8018
worse	0.1018	0.2124	0.3854	0.5996	0.8060
Worse	0.1060	0.2148	0.3790	0.6099	0.8180

80% power at HR = 0.6

Trial 2 – Single Arm, assume control rate is 0.03

- All patients on treatment arm.
- Assume control rate 0.03 from historical data (current patient population may not match!)

YIKES! →

	1.00	0.90	0.80	0.70	0.60
Better	0.6278	0.8858	0.9833	0.9992	1.0000
better	0.3116	0.6603	0.9168	0.9912	0.9997
Same	0.1063	0.3686	0.7482	0.9601	0.9977
worse	0.0245	0.1518	0.4955	0.8678	0.9894
Worse	0.0044	0.0463	0.2593	0.6904	0.9601

Power usually higher

Trial 3 – Heuristic informative prior based on historical data

- 25 patients on control, 75 on treatment
- Prior on control rate $\text{Gamma}(20, 20 \times 0.03)$
(this prior places mass between 0.02 and 0.04)

	1.00	0.90	0.80	0.70	0.60	Power still higher than 50-50
Better	0.1604	0.3340	0.5684	0.8019	0.9431	
better	0.1012	0.2479	0.4800	0.7372	0.9202	
Same	0.0606	0.1753	0.3958	0.6735	0.8893	
worse	0.0362	0.1174	0.3109	0.5956	0.8556	
Worse	0.0200	0.0770	0.2433	0.5172	0.8105	

lower type I error inflation

Bayesian Augmented Control

- In the previous trial, we informally guessed at a prior using the historical data. In Bayesian Augmented Control, we use a formal model to determine the prior.
- The amount of borrowing of information is determined by a hierarchical model incorporating the current and historical control arms.
- The amount of borrowing depends on the data!

Hierarchical Modeling

- Intuition – if one sees consistent (NOT necessarily identical) results in several historical trials, that provides information useful for predicting the current trial.
- Important – the data are used to determine similarity, no “structural” requirement the current control is similar to historical data. **The data can indicate similarity or differences between the current or historical data. The model adjusts.**

Model (Continuous, not TTE)

- Current control group $Y_0 \sim N(\gamma_0, sd)$ with historical trials $Y_1, \dots, Y_H$ each $Y_h \sim N(\gamma_h, sd)$. Means differ in each trial. (Data level of hierarchy)
- Hierarchical model (across trials distribution)
 $\gamma_0, \gamma_1, \dots, \gamma_H \sim N(\mu, \tau)$.
- Priors $\pi(\mu) = N(M, S)$ and $\pi(\tau) = \text{IGamma}(\alpha, \beta)$
- τ controls similarly between trials, and the data (current and historical) is used to estimate τ

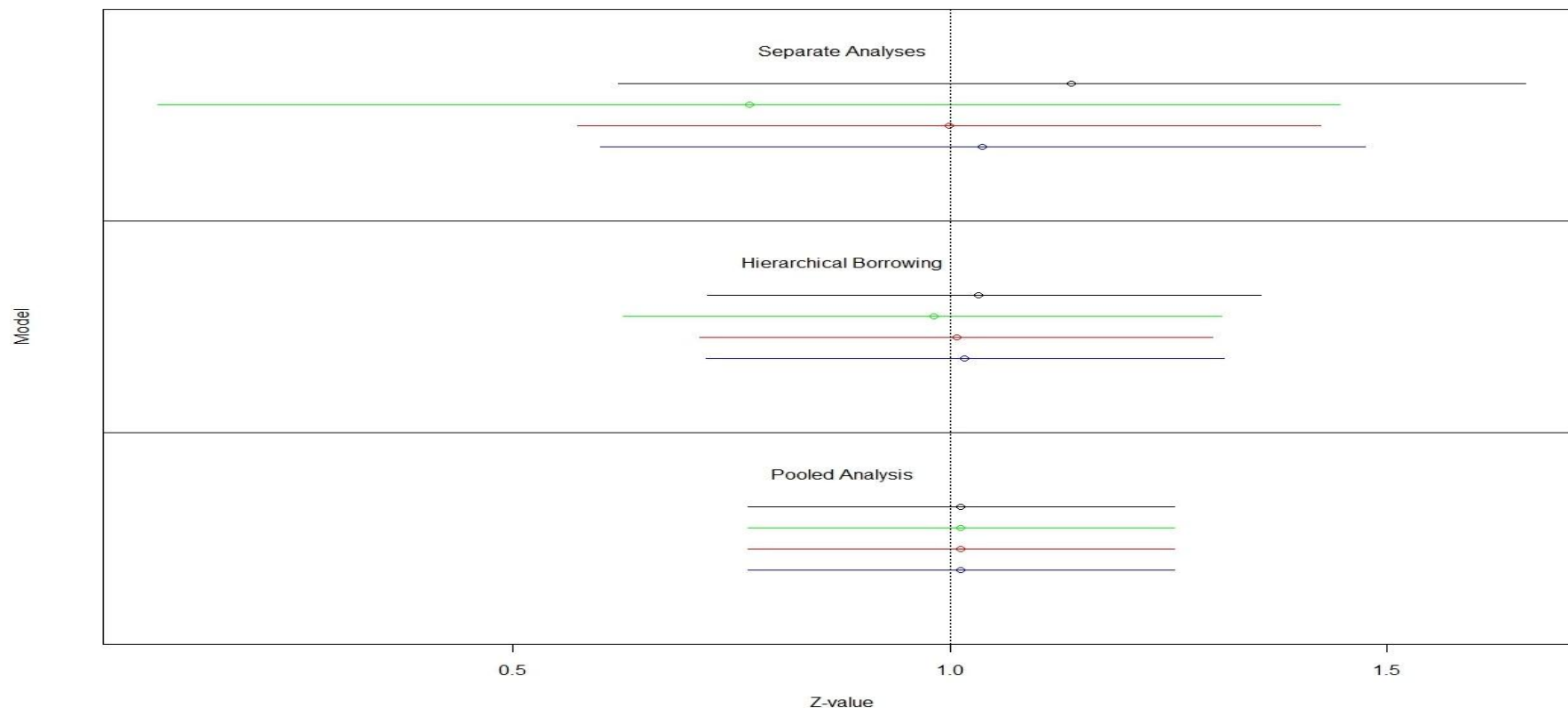
Interpreting $\gamma_0, \gamma_1, \dots, \gamma_H \sim N(\mu, \tau)$

- The across group mean μ provides an average of the control arm across trials.
- The standard deviation τ is vital to the model. Small τ indicate the trials have similar mean, while large τ indicate major differences between trials.
- τ is estimated from the data (not pre-specified)
- τ near 0 results in near pooling of the data
- τ large results in minimal borrowing (little common information)
- When appropriate, borrowing produces more accurate statistical results. The hierarchical model automatically assesses the appropriate amount of borrowing and uses that to produce more accurate estimates.

Interpreting $\gamma_0, \gamma_1, \dots, \gamma_H \sim N(\mu, \tau)$

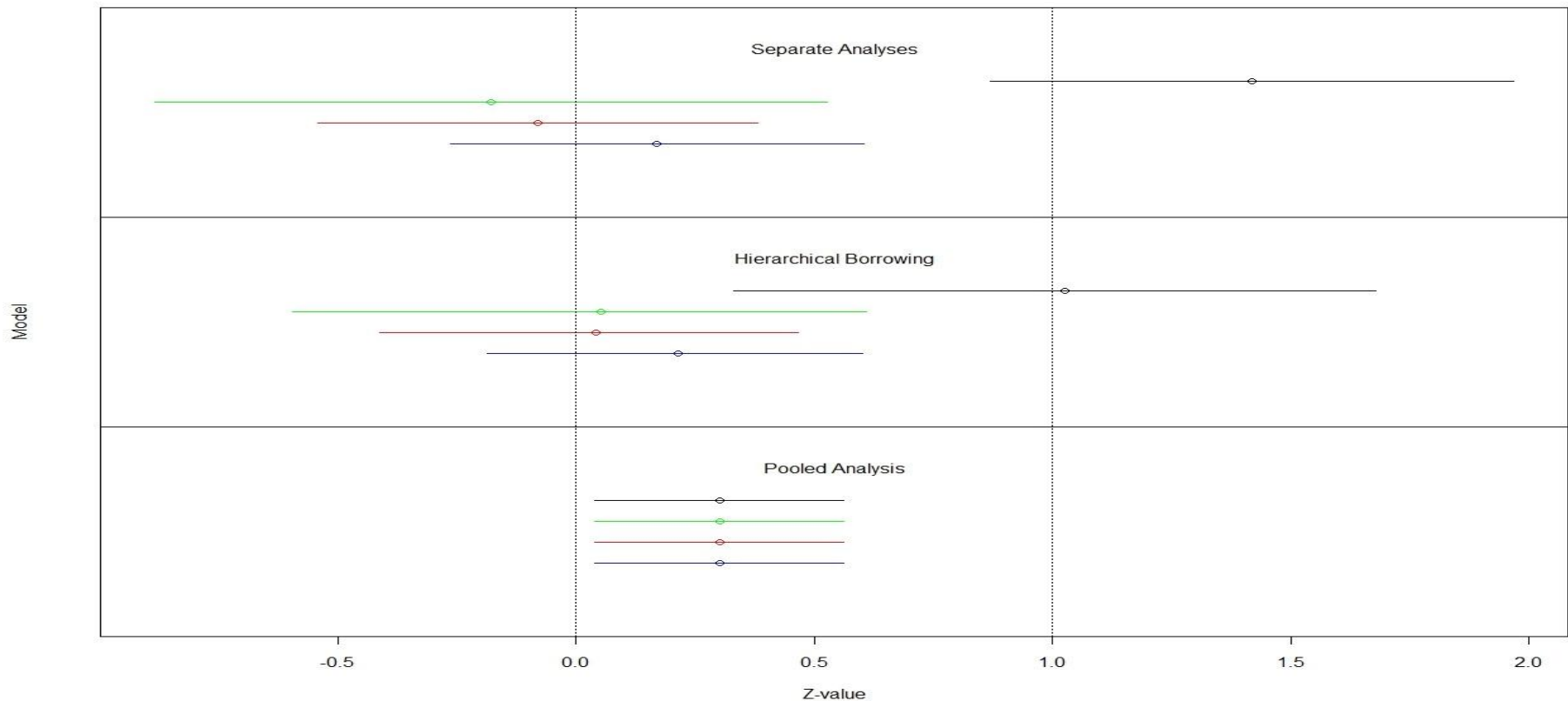
- We show the information borrowing for two situations – when the current control group appears similar to previous data and when the current control group appears different.
- We contrast three approaches
 - Separate analyses – no information borrowed from historical data
 - BAC (Hierarchical Modeling)
 - Pooling – historical control subjects pooled with current control subjects

Consistent Control Groups (current control in black)



Here x-axis is showing z-scores (standardized treatment effects). The borrowing “shrinks” the results closer together as we move from separate models to a pooled model. In this case the hierarchical model is similar to pooling.

Current control (black) different



Borrowing is reduced for this model based on a dissimilarity of the current control group. However, certainly the large separate effect of the current control group is pulled down by the historical data, but NOT as much as pooling.

TTE BAC is slightly more involved

- Hierarchical Borrowing occurs on the log hazard ratios
- $\lambda_h = \lambda \exp(\gamma_h)$
- $\gamma_0, \gamma_1, \dots, \gamma_H \sim N(\mu, \tau)$ with hyperpriors as before.

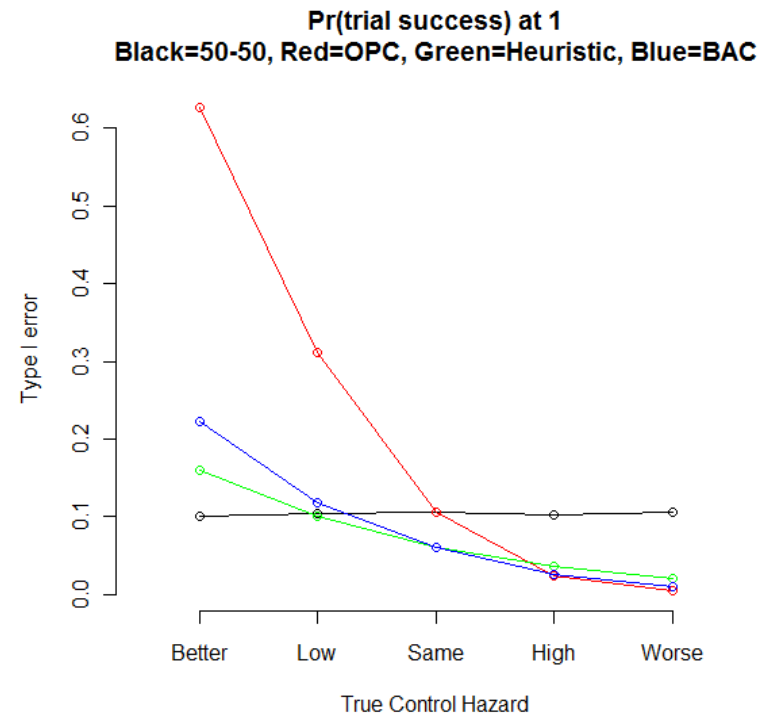
Trial 4 – formal BAC

- 25 patients on control, 75 on treatment
- BAC, $\mu \sim N(0,3)$ and $\tau \sim \text{IGamma}(1,0.06)$

	1.00	0.90	0.80	0.70	0.60
Better	0.2228	0.4511	0.7180	0.9031	0.9842
better	0.1186	0.3074	0.5869	0.8429	0.9710
Same	0.0598	0.1844	0.4375	0.7542	0.9407
worse	0.0258	0.1108	0.3146	0.6541	0.9047
Worse	0.0109	0.0595	0.2145	0.5307	0.8424

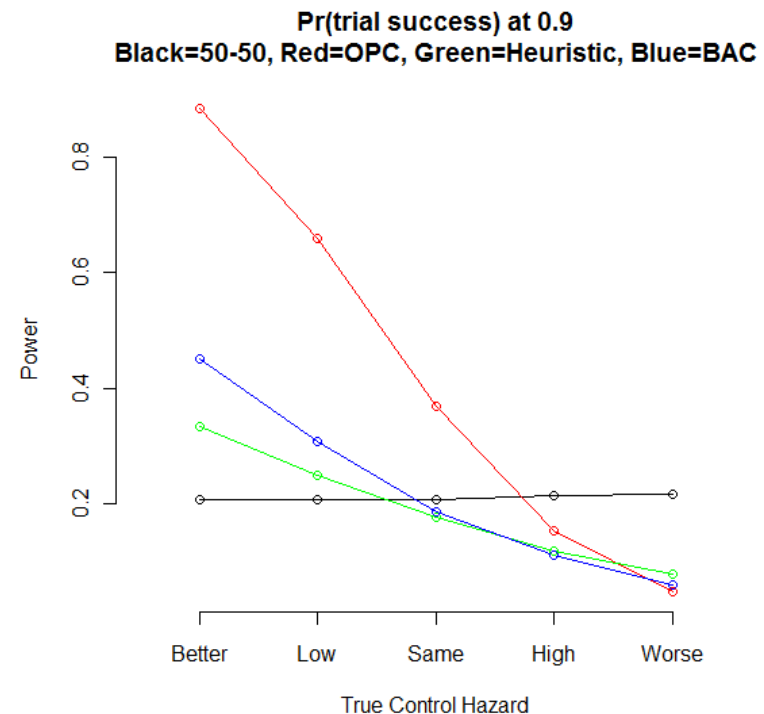
Type I error comparison

- OPC has excessively high type I error
- BAC and Heuristic do not inflate Type I error when current controls are in the range of historical controls.



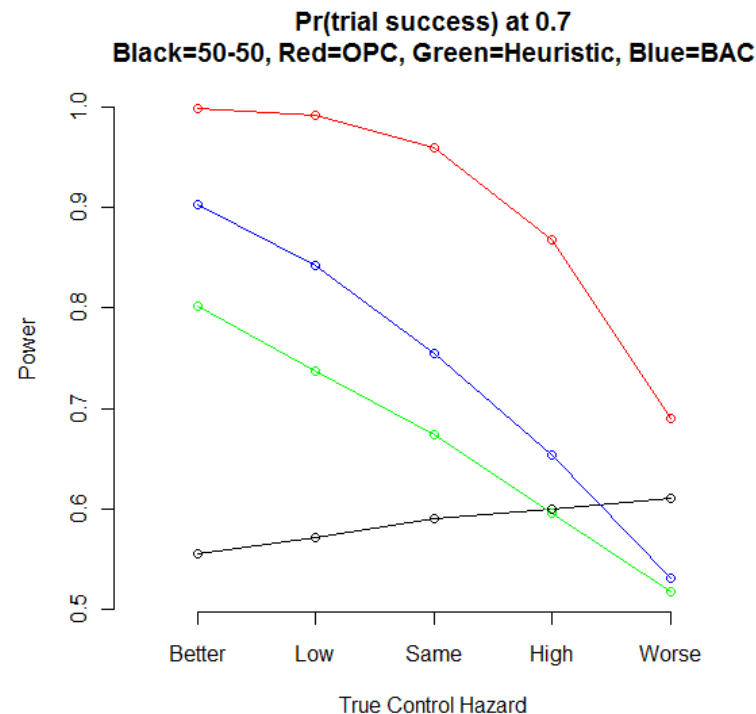
Power comparison at HR=0.9

- Power is weak for all except the “disqualified” OPC.
- Power higher for Heuristic and BAC (compared to 50-50) when population is better.



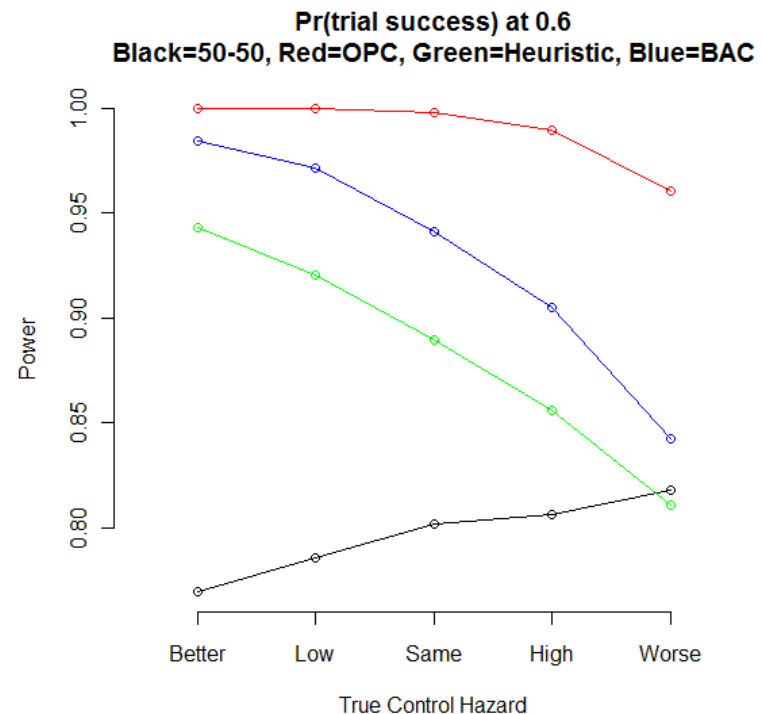
Power comparison at HR=0.7

- BAC has large (10-20%) power gains over 50-50 when population is in the range of historical controls or better.
- BAC reduces power slightly when population is significantly worse.



Power comparison at HR=0.6

- Increased sample size now pays off for all considered control hazards, with BAC uniformly more powerful.
- Power gains are significant for BAC over 50-50.



Conclusions

- Bayesian Augmented Control provides a mechanism for incorporating historical information into an existing trial.
- BAC increases power in the majority of cases (important to understand where it doesn't)
- Type I error inflation is mild/moderate unless current control population is significantly better than historical controls.