Bayesian Adjustment for Multiplicity

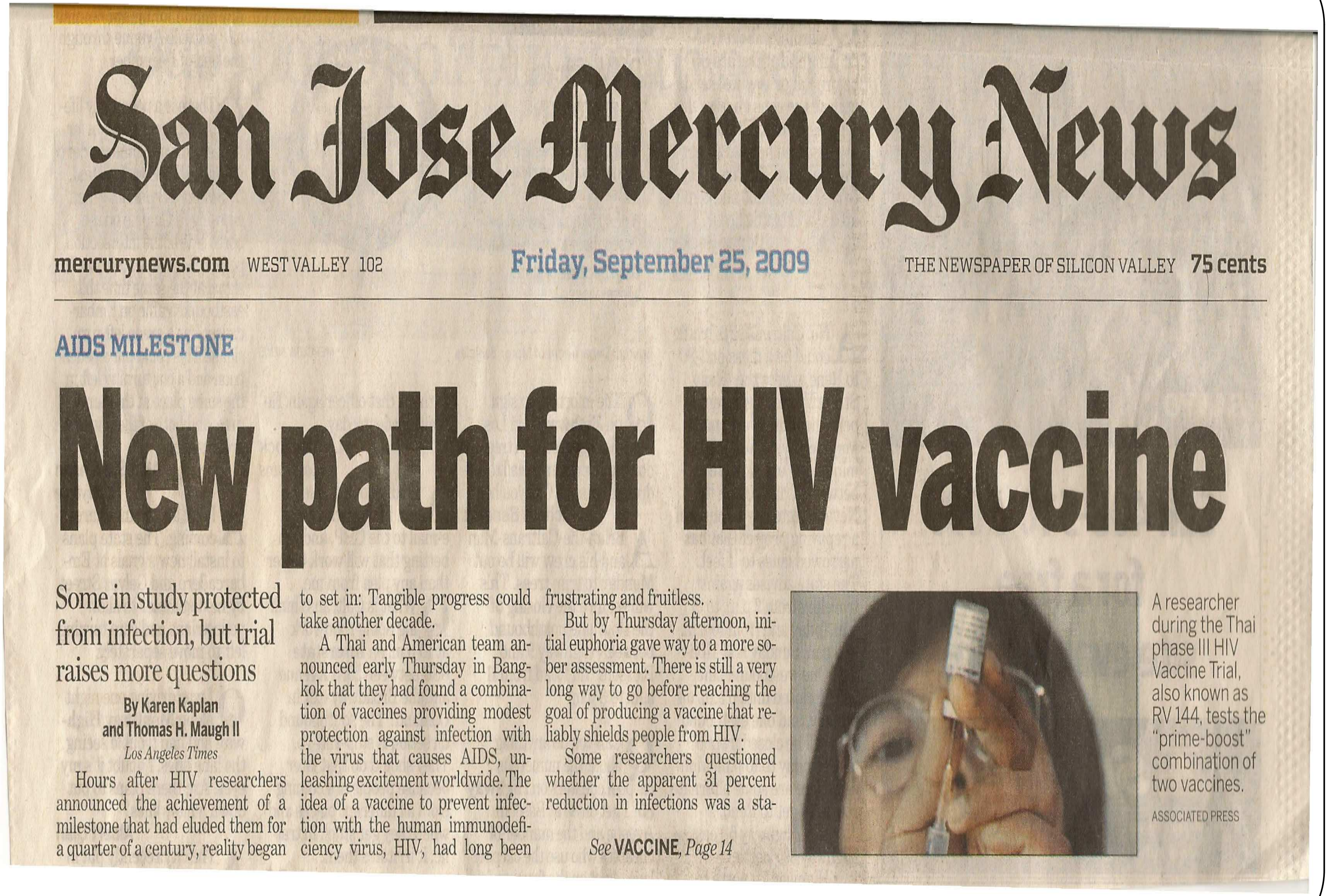
Jim Berger

Duke University
Statistical and Applied Mathematical Sciences Institute

with James Scott

University of Texas

Subjective Bayes 2009
December 14-16, 2009



San Jose Mercury News

mercurynews.com WEST VALLEY 102 Friday, September 25, 2009 THE NEWSPAPER OF SILICON VALLEY 75 cents

AIDS MILESTONE

New path for HIV vaccine

Some in study protected from infection, but trial raises more questions

By Karen Kaplan and Thomas H. Maugh II

Los Angeles Times

Hours after HIV researchers announced the achievement of a milestone that had eluded them for a quarter of a century, reality began to set in: Tangible progress could take another decade.

A Thai and American team announced early Thursday in Bangkok that they had found a combination of vaccines providing modest protection against infection with the virus that causes AIDS, unleashing excitement worldwide. The idea of a vaccine to prevent infection with the human immunodeficiency virus, HIV, had long been frustrating and fruitless.

But by Thursday afternoon, initial euphoria gave way to a more sober assessment. There is still a very long way to go before reaching the goal of producing a vaccine that reliably shields people from HIV.

Some researchers questioned whether the apparent 31 percent reduction in infections was a sta-

See VACCINE, Page 14

A researcher during the Thai phase III HIV Vaccine Trial, also known as RV 144, tests the "prime-boost" combination of two vaccines.

ASSOCIATED PRESS

Hypotheses and Data:

- Alvac had shown no effect
- Aidsvax had shown no effect

Question: Would Alvac as a primer and Aidsvax as a booster work?

The Study: Conducted in Thailand with 16,395 individuals from the general (not high-risk) population:

- 71 HIV cases reported in the 8198 individuals receiving placebos
- 51 HIV cases reported in the 8197 individuals receiving the treatment

The test that was likely performed:

- Let p_1 and p_2 denote the probability of HIV in the placebo and treatment populations, respectively.
- Test $H_0 : p_1 = p_2$ versus $H_1 : p_1 > p_2$
  (vaccines were not live, so $p_1 < p_2$ can probably be ignored)
- Normal approximation okay, so

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sigma_{\{\hat{p}_1 - \hat{p}_2\}}} = \frac{.00866 - .00622}{.00134} = 1.82$$

  is approximately $\mathrm{N}(\theta, 1)$, where $\theta = (p_1 - p_2)/(.00134)$.
  We thus test $H_0 : \theta = 0$ versus $H_1 : \theta > 0$, based on z.
- Observed $z = 1.82$, so the (one-sided) p-value is 0.034.

Bayesian Analysis:

Prior distribution:

- $Pr(H_i)$ = prior probability that H_i is true, $i = 0, 1$,
- On $H_1 : \theta > 0$, let $\pi(\theta)$ be the prior density for θ.

Note: H_0 must be believable (at least approximately) for this to be reasonable (i.e., no fake nulls).

Subjective Bayes: choose these based on personal beliefs

Objective (or default) Bayes: choose

- $Pr(H_0) = Pr(H_1) = \frac{1}{2}$,
- $\pi(\theta) = \text{Uniform}(0, 6.46)$, which arises from assigning
  - uniform for p_2 on $0 < p_2 < p_1$,
  - plug in for p_1 .

Posterior probability of hypotheses:

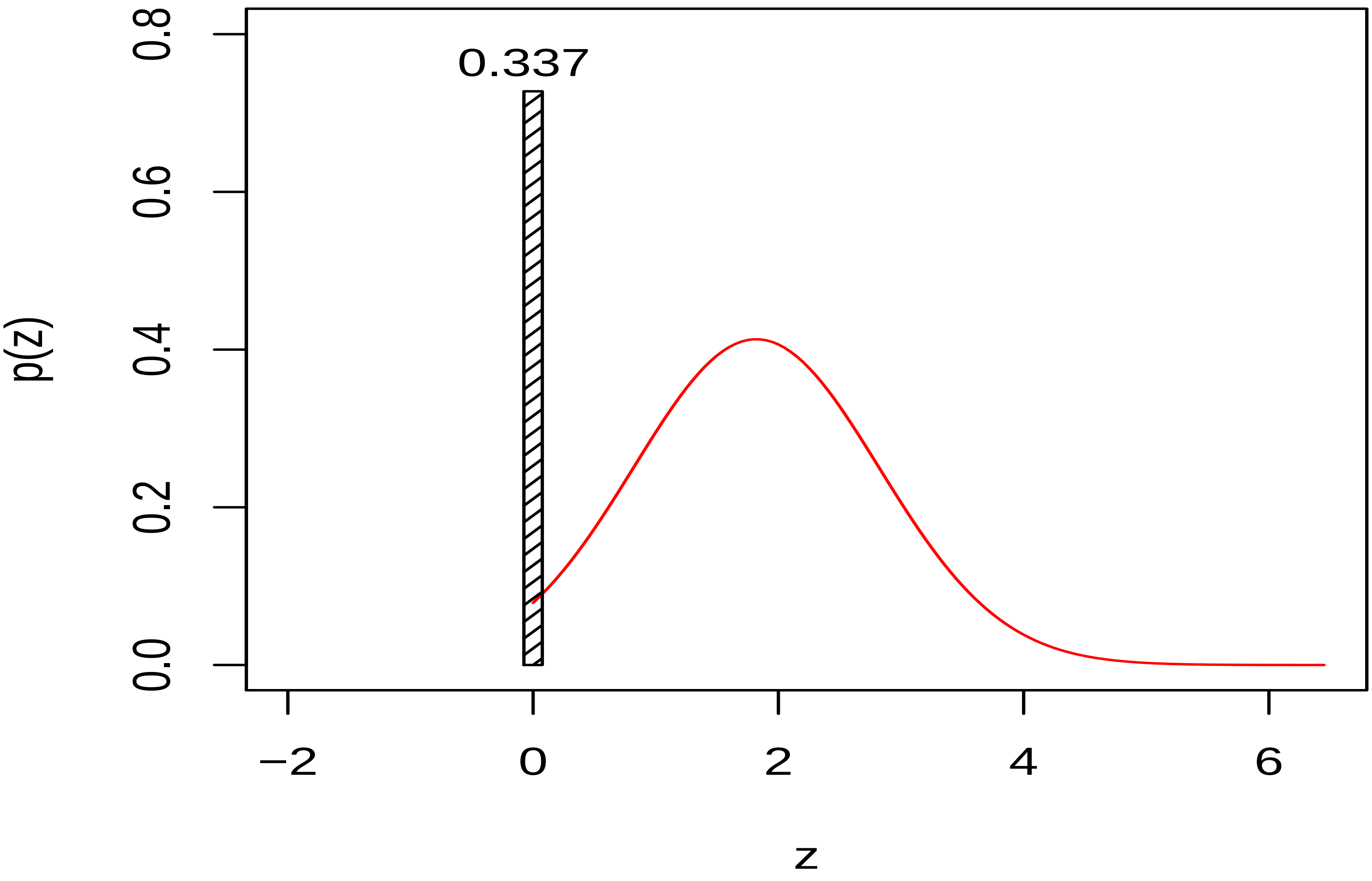
$$\begin{aligned} Pr(H_0|z) &= \text{probability that } H_0 \text{ true, given data } z \\ &= \frac{f(z\,|\,\theta=0)\,Pr(H_0)}{Pr(H_0)\,f(x\,|\,\theta=0)+Pr(H_1)\int_0^\infty f(z\,|\,\theta)\pi(\theta)d\theta} \end{aligned}$$

For the objective prior, $Pr(H_0 \mid z=1.82) \approx 0.337$ (recall, p-value $\approx .034$)

Posterior density on $H_1 : \theta > 0$ is

$$\pi(\theta|z=1.82, H_1) \propto \pi(\theta) f(1.82 \mid \theta) = (0.413)e^{-\frac{1}{2}(1.82-\theta)^2}$$

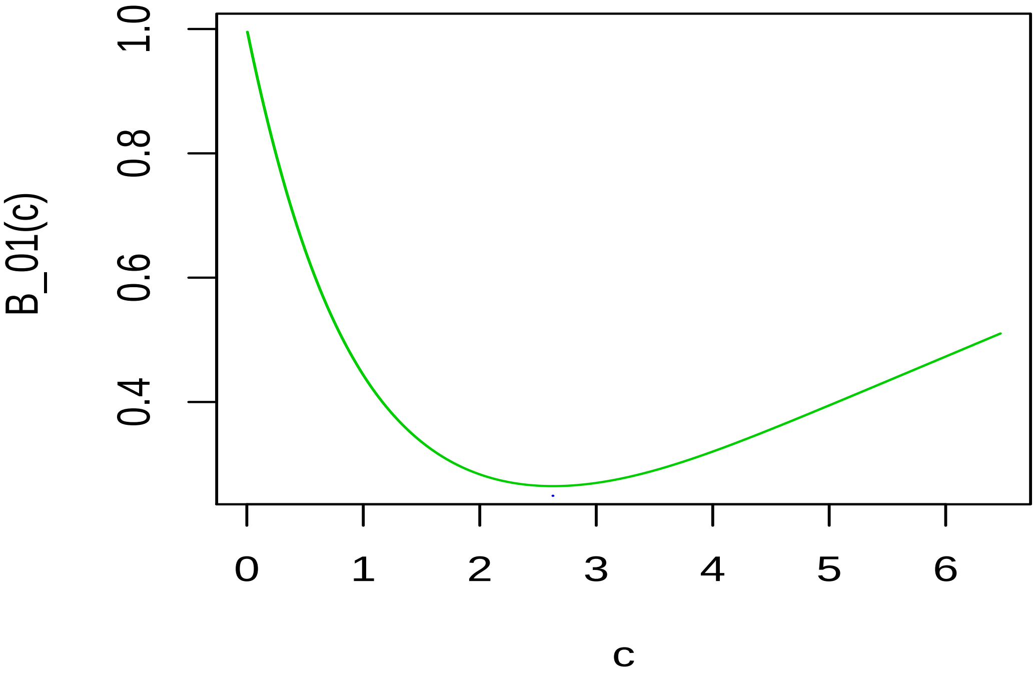
for $0 < \theta < 6.46$.


0.337
p(z)
0.0
0.2
0.4
0.6
0.8
−2
0
2
4
6
z

Robust Bayes: Report the *Bayes factor* (the odds of H_0 to H_1) as a function of $\pi_C(\theta) \equiv \text{Uniform}(0, C)$:

$$B_{01}(C) = \frac{\text{likelihood of H}_0\text{ for observed data}}{\text{average likelihood of H}_1} = \frac{\frac{1}{\sqrt{2\pi}}e^{-(1.82)^2/2}}{\int_0^C \frac{1}{\sqrt{2\pi}}e^{-(1.82-\theta)^2/2}C^{-1}d\theta}$$

.



Note: $\min_C B_{01}(C) = 0.265$ (while $B_{01}(6.46) = 0.51$).

Note: This is the same Bayes factor envelope for nonincreasing priors.

Outline

- Background on multiplicity
- Bayesian approach to control of multiplicity
- A simple example: multiple testing under exclusivity
- Variable selection (including comparison with empirical Bayes)
- Subgroup analysis

Multiplicity Arising in SAMSI Programs

- Stochastic Computation / Data Mining and Machine Learning
  - *Example:* Analysis of gene expression microarrays, with tests concerning the mean differential expression, μ_i, of genes $i = 1, \ldots, 10,000$:

    $$H_0 : \mu_i = 0 \quad \text{versus} \quad H_1 : \mu_i \neq 0\,.$$

    Multiplicity problem: Even if all $\mu_i = 0$, one would find that roughly 500 tests reject at, say, level $\alpha = 0.05$, so a correction for this effect is needed.
- National Defense and Homeland Security
  - *Example:* In Syndromic Surveillance, many counties in the USA perform daily tests on the 'excess' of some symptoms, the goal being early detection of the outbreak of epidemics or of bio-terrorist attacks.
- Astrostatistics
  - *Example:* 1.6 million tests of CMB radiation for non-Gaussianity in the spatial distribution.
- Latent Variable Models in the Social Sciences
  - *Example:* Variable selection in structural equation modeling

- Multiplicity and Reproducibility in Scientific Studies

  Additional motivations for the program:

  - Multiplicity adjustment is often ignored, because of
    - lack of understanding of the importance of the issue
      - *American Scientist* (January 2007) article about personalized medicine barely mentioned the problem.
      - *Nature* (January 2007) article reviewing the status of personalized medicine didn't mention multiplicity at all.
    - the lack of suitable adjustment methodology
  - Indications of an increasing problem with reproducibility in science
    - In the USA, drug compounds entering Phase I development today have an 8% chance of reaching market, versus a 14% chance 15 years ago
    - Even our most rigorously controlled statistical analyses do not seem to be immune:
      - 50% phase III failure rates are now being reported, versus a 20% failure rate 10 years ago
      - reports that 30% of phase III successes fail to replicate

General Approach to Bayesian Multiplicity Adjustment

1. Represent the problem as a *model uncertainty* problem: Models $\mathcal{M}_i$, with densities $f_i(\mathbf{x} \mid \boldsymbol{\theta}_i)$ for data $\mathbf{x}$, given unknown parameters $\boldsymbol{\theta}_i$; prior distributions $\pi_i(\boldsymbol{\theta}_i)$; and marginal likelihoods $m_i(\mathbf{x}) = \int f_i(\mathbf{x} \mid \boldsymbol{\theta}_i)\pi_i(\boldsymbol{\theta}_i)d\boldsymbol{\theta}_i$.

2. Specify prior probabilities, $P(\mathcal{M}_i)$, of models to reflect the multiplicity issues; **Bayesian analysis controls multiplicity through** $P(\mathcal{M}_i)$ [a]

- *Subjective Bayesian Analysis:* If the $P(\mathcal{M}_i)$ are real subjective probabilities, or arise from subjective modeling of the probabilities, that's it: multiplicity correction has been done.
- *Objective Bayesian Analysis:* One has to be careful to make choices of the $P(\mathcal{M}_i)$ that ensure multiplicity correction (e.g., specifying equal prior probabilities does *not* generally control multiplicity)!

3. Implement Bayesian model averaging (model selection?), based on

$$P(\mathcal{M}_i \mid \mathbf{x}) = \frac{P(\mathcal{M}_i)\ m_i(\mathbf{x})}{\sum_{j=1}^{k} P(\mathcal{M}_j)\ m_j(\mathbf{x})}\ .$$

[a] see, e.g., Jeffreys 1961; Waller and Duncan 1969; Meng and Demptster 1987; Berry 1988; Westfall, Johnson and Utts 1997; Carlin and Louis 2000.

Simple Example: Multiple Testing under Exclusivity

Suppose one is testing mutually exclusive hypotheses $H_i,\ i = 1, \ldots, m$, so each hypothesis is a separate model. If the hypotheses are viewed as exchangeable, choose $P(H_i) = P(\mathcal{M}_i) = 1/m$.

Example: 1000 energy channels (or 10^{12} at CERN) are searched for a signal:

- if the signal is known to exist and occupy only one channel, but no channel is theoretically preferred, each channel can be assigned prior probability 0.001.
- if the signal is not known to exist (e.g., it is the prediction of a non-standard physics theory) prior probability 1/2 should be given to 'no signal,' and probability 0.0005 to each channel.

Note: this is the answer regardless of the data structure.

Note: equal prior model probabilities does provide multiplicity control here.

Variable Selection

Problem: Data $\mathbf{X}$ arises from a normal linear regression model, with m possible regressors having associated unknown regression coefficients $\beta_i, i = 1, \ldots m$, and unknown variance σ^2.

Models: Consider selection from among the submodels $\mathcal{M}_i$, $i = 1, \ldots, 2^m$, having only k_i regressors with coefficients $\boldsymbol{\beta}_i$ (a subset of $(\beta_1, \ldots, \beta_m)$) and resulting density $f_i(\mathbf{x} \mid \boldsymbol{\beta}_i, \sigma^2)$.

Prior density under $\mathcal{M}_i$: Zellner-Siow priors $\pi_i(\boldsymbol{\beta}_i, \sigma^2)$ in examples.

Marginal likelihood of $\mathcal{M}_i$: $m_i(\mathbf{x}) = \int f_i(\mathbf{x} \mid \boldsymbol{\beta}_i, \sigma^2)\pi_i(\boldsymbol{\beta}_i, \sigma^2)\, d\boldsymbol{\beta}_i d\sigma^2$

Prior probability of $\mathcal{M}_i$: $P(\mathcal{M}_i)$

Posterior probability of $\mathcal{M}_i$:

$$P(\mathcal{M}_i \mid \mathbf{x}) = \frac{P(\mathcal{M}_i)m_i(\mathbf{x})}{\sum_j P(\mathcal{M}_j)m_j(\mathbf{x})} \,.$$

Common Choices of the $P(\mathcal{M}_i)$

Equal prior probabilities: $P(\mathcal{M}_i) = 2^{-m}$

Bayes exchangeable variable inclusion:

- Each variable, β_i, is independently in the model with unknown probability p (called the prior inclusion probability).
- p has a Beta$(p \mid a, b)$ distribution, chosen to represent prior beliefs concerning the unknown p.
- Then, since k_i is the number of variables in model $\mathcal{M}_i$,

$$P(\mathcal{M}_i) = \int_0^1 p^{k_i}(1-p)^{m-k_i}\text{Beta}(p \mid a, b)dp = \frac{Beta(a+k_i, b+m-k_i)}{Beta(a,b)}\ .$$

Empirical Bayes exchangeable variable inclusion: Find the MLE $\hat{p}$ by maximizing the marginal likelihood of p, $\sum_j p^{k_j}(1-p)^{m-k_j} m_j(\mathbf{x})$, and use $P(\mathcal{M}_i) = \hat{p}^{k_i}(1-\hat{p})^{m-k_i}$ as the prior model probabilities.

Controlling for multiplicity in variable selection

Equal prior probabilities: $P(\mathcal{M}_i) = 2^{-m}$ does *not* control for multiplicity here; it corresponds to fixed prior inclusion probability $p = 1/2$ for each variable, which is rarely appropriate. (Ley and Steel (2007) show other inadequacies of this choice.)

Empirical Bayes exchangeable variable inclusion does control for multiplicity, in that $\hat{p}$ will be small if there are many β_i that are zero.

Bayes exchangeable variable inclusion also controls for multiplicity (see Scott and Berger, 2008), although the $P(\mathcal{M}_i)$ are fixed.

Note: The control of multiplicity by Bayes and EB variable inclusion usually reduces model complexity, but is *different* than the usual Bayesian Ockham's razor effect that reduces model complexity.

- The Bayesian Ockham's razor operates through the effect of model priors $\pi_i(\boldsymbol{\beta}_i, \sigma^2)$ on $m_i(\mathbf{x})$, penalizing models with more parameters.
- Multiplicity correction occurs through the choice of the $P(\mathcal{M}_i)$.

	Equal model probabilities				**Bayes variable inclusion**			
	Number of noise variables				Number of noise variables			
Signal	1	10	40	90	1	10	40	90
$\beta_1 : -1.08$	.999	.999	.999	.999	.999	.999	.999	.999
$\beta_2 : -0.84$	.999	.999	.999	.999	.999	.999	.999	.988
$\beta_3 : -0.74$	.999	.999	.999	.999	.999	.999	.999	.998
$\beta_4 : -0.51$	.977	.977	.999	.999	.991	.948	.710	.345
$\beta_5 : -0.30$	.292	.289	.288	.127	.552	.248	.041	.008
$\beta_6 : +0.07$	.259	.286	.055	.008	.519	.251	.039	.011
$\beta_7 : +0.18$	.219	.248	.244	.275	.455	.216	.033	.009
$\beta_8 : +0.35$	.773	.771	.994	.999	.896	.686	.307	.057
$\beta_9 : +0.41$	.927	.912	.999	.999	.969	.861	.567	.222
$\beta_{10} : +0.63$	.995	.995	.999	.999	.996	.990	.921	.734
False Positives	0	2	5	10	0	1	0	0

Table 1: Posterior inclusion probabilities for 10 real variables in a simulated data set, with a uniform prior chosen for p.

Comparison of Bayes and Empirical Bayes Approaches

Theorem 1 *In the variable-selection problem, if the null model (or full model) has the largest marginal likelihood,* $m(\mathbf{x})$*, among all models, then the MLE of* p *is* $\hat{p} = 0$ *(or* $\hat{p} = 1$*.)* (The naive EB approach, which assigns $P(\mathcal{M}_i) = \hat{p}^{k_i}(1-\hat{p})^{m-k_i}$, concludes that the null (full) model has probability 1.)

A simulation with 10,000 repetitions to gauge the severity of the problem:

- $m = 14$ covariates, orthogonal design matrix
- p drawn from $U(0,1)$; regression coefficients are 0 with probability p and drawn from a Zellner-Siow prior with probability $(1-p)$.
- $n = 16$, 60, and 120 observations drawn from the given regression model.

Case	$\hat{p} = 0$	$\hat{p} = 1$
$n = 16$	820	781
$n = 60$	783	766
$n = 120$	723	747

Is empirical Bayes at least accurate asymptotically as $m \to \infty$?

Posterior model probabilities, given p:

$$P(\mathcal{M}_i \mid \mathbf{x}, p) = \frac{p^{k_i}(1-p)^{m-k_i} m_i(\mathbf{x})}{\sum_j p^{k_j}(1-p)^{m-k_j} m_j(\mathbf{x})}$$

Posterior distribution of p: $\pi(p \mid \mathbf{x}) = K \sum_j p^{k_j}(1-p)^{m-k_j} m_j(\mathbf{x})$

This *does* concentrate about the true p as $m \to \infty$, so one might expect that

$$P(\mathcal{M}_i \mid \mathbf{x}) = \int_0^1 P(\mathcal{M}_i \mid \mathbf{x}, p)\pi(p \mid \mathbf{x}) dp \approx P(\mathcal{M}_i \mid \mathbf{x}, \hat{p}) \propto m_i(\mathbf{x})\ \hat{p}^{k_i}(1-\hat{p})^{m-k_i}.$$

This is not necessarily true; indeed

$$\begin{aligned} \int_0^1 P(\mathcal{M}_i \mid \mathbf{x}, p)\pi(p \mid \mathbf{x}) dp &= \int_0^1 \frac{p^{k_i}(1-p)^{m-k_i} m_i(\mathbf{x})}{\pi(p \mid \mathbf{x})/K} \times \pi(p \mid \mathbf{x})\ dp \\ &\propto m_i(\mathbf{x}) \int_0^1 p^{k_i}(1-p)^{m-k_i}\pi(p) dp \propto m_i(\mathbf{x}) P(\mathcal{M}_i)\,. \end{aligned}$$

Caveat: Some EB techniques have been justified; see Efron and Tibshirani (2001), Johnstone and Silverman (2004), Cui and George (2006), and Bogdan et. al. (2008).

Theorem 2 *Suppose the true model size k_T satisfies $k_T/m = p_T + O(1/\sqrt{m})$ as $m \to \infty$, where $0 < p_T < 1$. Consider all models M_i such that $k_T - k_i = O(\sqrt{m})$, and consider the optimal situation for EB in which*

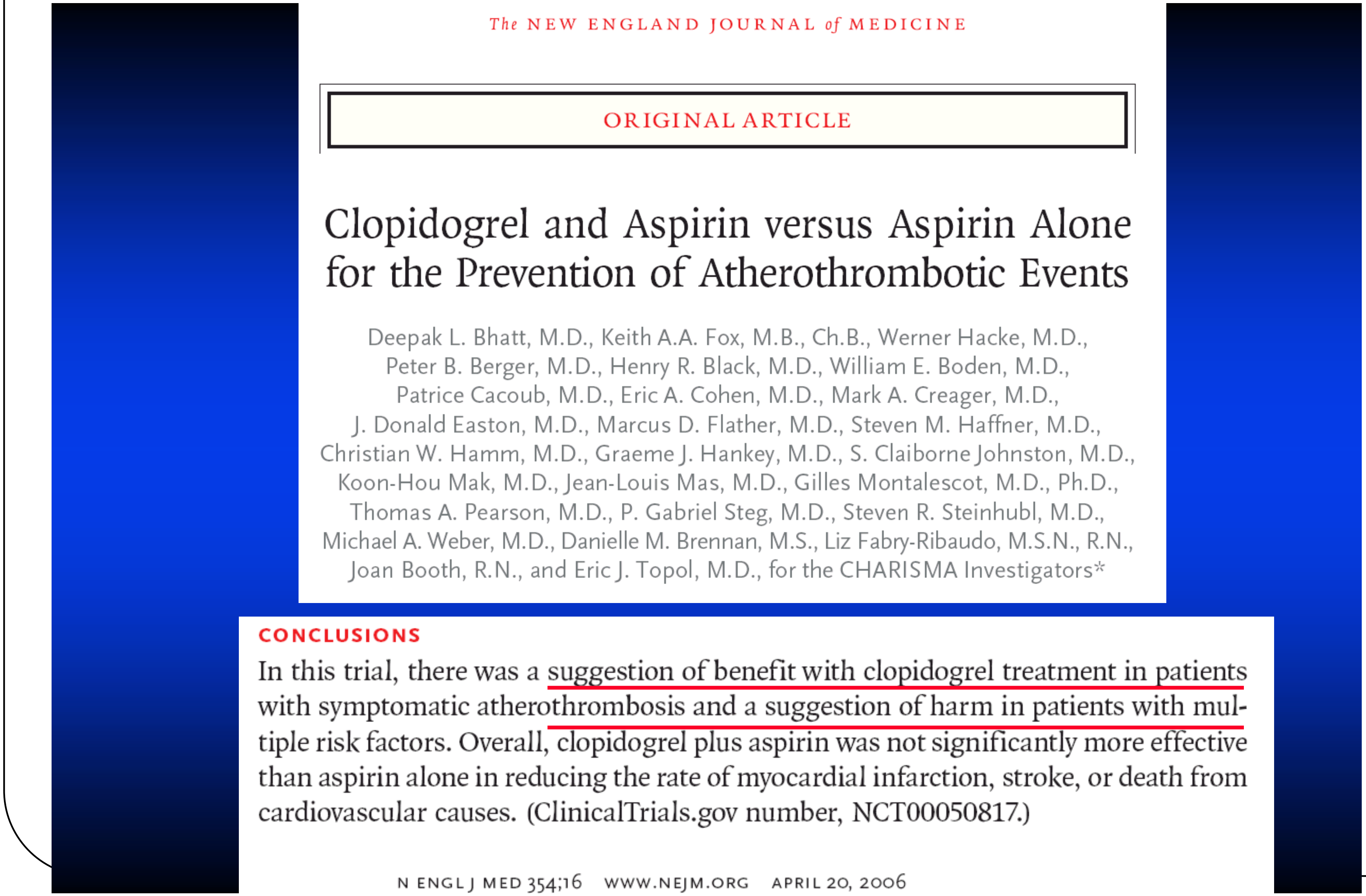
$$\hat{p} = p_T + O(\frac{1}{\sqrt{m}}) \quad as \quad m \to \infty\,.$$

Then the ratio of the prior probabilities assigned to such models by the Bayes approach and the empirical Bayes approach satisfies

$$\frac{P_B(\mathcal{M}_i)}{P_{EB}(\mathcal{M}_i)} = \frac{\int_0^1 p^{k_i}(1-p)^{m-k_i}\pi(p)dp}{(\hat{p})^{k_i}(1-\hat{p})^{m-k_i}} = O\left(\frac{1}{\sqrt{m}}\right)\,,$$

providing $\pi(\cdot)$ is continuous and nonzero.

Subgroup Analysis



The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

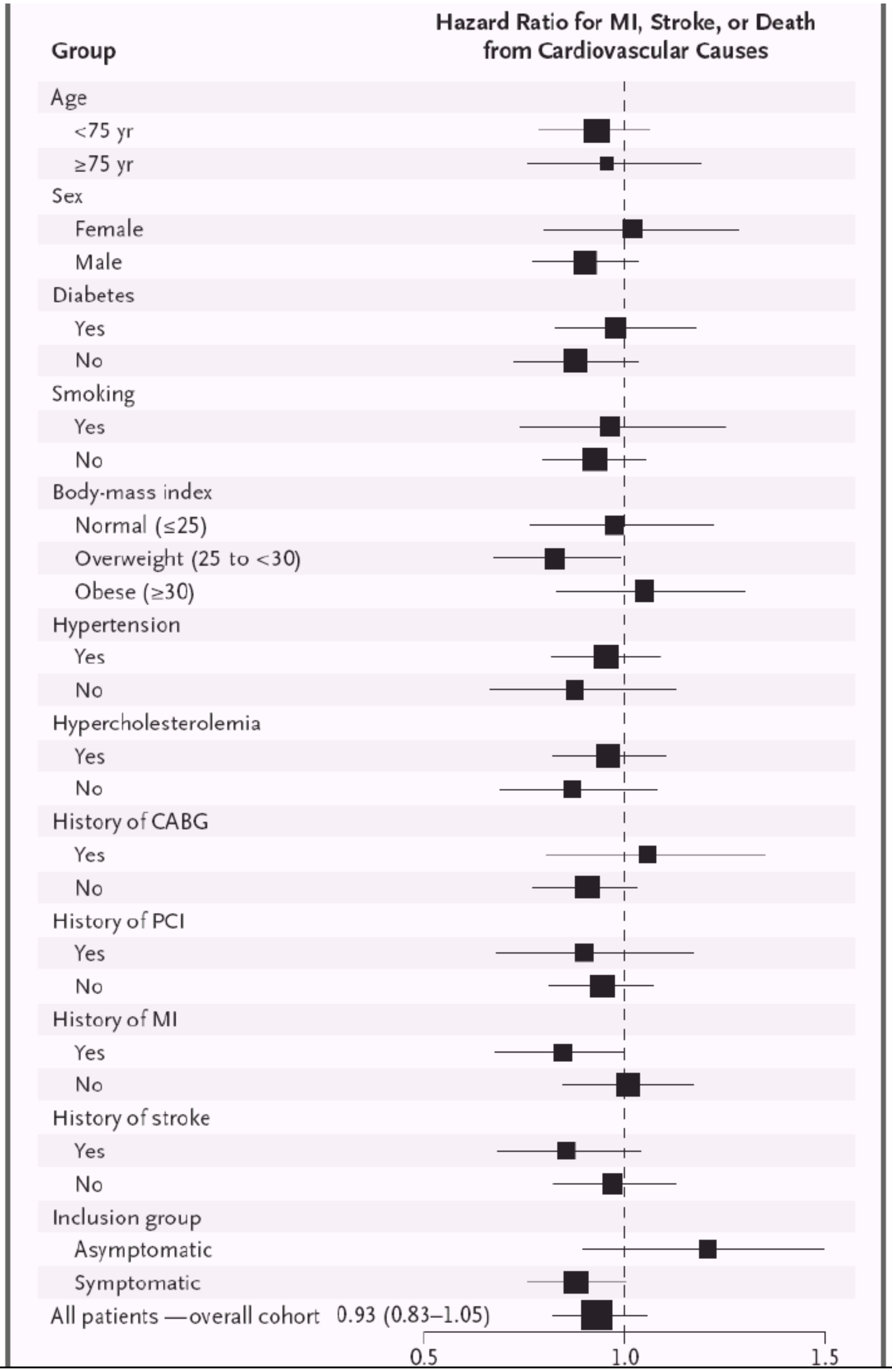
Clopidogrel and Aspirin versus Aspirin Alone for the Prevention of Atherothrombotic Events

Deepak L. Bhatt, M.D., Keith A.A. Fox, M.B., Ch.B., Werner Hacke, M.D., Peter B. Berger, M.D., Henry R. Black, M.D., William E. Boden, M.D., Patrice Cacoub, M.D., Eric A. Cohen, M.D., Mark A. Creager, M.D., J. Donald Easton, M.D., Marcus D. Flather, M.D., Steven M. Haffner, M.D., Christian W. Hamm, M.D., Graeme J. Hankey, M.D., S. Claiborne Johnston, M.D., Koon-Hou Mak, M.D., Jean-Louis Mas, M.D., Gilles Montalescot, M.D., Ph.D., Thomas A. Pearson, M.D., P. Gabriel Steg, M.D., Steven R. Steinhubl, M.D., Michael A. Weber, M.D., Danielle M. Brennan, M.S., Liz Fabry-Ribaudo, M.S.N., R.N., Joan Booth, R.N., and Eric J. Topol, M.D., for the CHARISMA Investigators*

CONCLUSIONS

In this trial, there was a suggestion of benefit with clopidogrel treatment in patients with symptomatic atherothrombosis and a suggestion of harm in patients with multiple risk factors. Overall, clopidogrel plus aspirin was not significantly more effective than aspirin alone in reducing the rate of myocardial infarction, stroke, or death from cardiovascular causes. (ClinicalTrials.gov number, NCT00050817.)

N ENGL J MED 354;16 WWW.NEJM.ORG APRIL 20, 2006


Group
Hazard Ratio for MI, Stroke, or Death from Cardiovascular Causes
Age
<75 yr
≥75 yr
Sex
Female
Male
Diabetes
Yes
No
Smoking
Yes
No
Body-mass index
Normal (≤25)
Overweight (25 to <30)
Obese (≥30)
Hypertension
Yes
No
Hypercholesterolemia
Yes
No
History of CABG
Yes
No
History of PCI
Yes
No
History of MI
Yes
No
History of stroke
Yes
No
Inclusion group
Asymptomatic
Symptomatic
All patients —overall cohort 0.93 (0.83–1.05)
0.5
1.0
1.5

Frequentist adjustment for performing 26 hypothesis tests

- Split the data into one part to suggest a subgroup and another part to confirm (or confirm with a new experiment).
- Bonferonni correction
  - To achieve an overall error probability level of 0.05 when conducting 26 tests, one would need to use a per-test rejection level of $\alpha = 0.05/26 = 0.002$.
  - This is likely much too conservative because of the dependence in the 26 tests.
- Various bootstrap types of correction to try to account for dependence.

Bayesian adjustment

Let $\boldsymbol{v}$ be the vector of 25 zeroes and ones indicating subgroup characteristics.

For each possible such vector, let $\mu_{\boldsymbol{v}}$ denote the mean of the intersected subgroup (e.g., young, male, diabetic, non-smoker,...).

Data: $\boldsymbol{x} \sim f(\boldsymbol{x} \mid \{\mu_{\boldsymbol{v}}, \text{all possible } \boldsymbol{v}\})$.

Two classes of approaches

- Factor-based approaches
- Aggregation-based approaches

An example factor-based approach

Model the intersected subgroup means additively as

$$\mu_{\boldsymbol{v}} = \mu + \boldsymbol{v}\boldsymbol{\beta}, \quad \boldsymbol{\beta} = (\beta_1, \ldots, \beta_{25})',$$

where μ is an overall mean and β_i is the effect corresponding to the i^{th} subgroup factor.

Conversion to model selection:

- Let $\boldsymbol{\gamma} = (\gamma_0, \boldsymbol{\gamma}^*) = (\gamma_0, \gamma_1, \ldots, \gamma_{25})$ be the vector of zeroes and ones, indicating whether μ (corresponding to γ_0) and each factor β_i is zero or not.
- This defines the model $\mathcal{M}_{\boldsymbol{\gamma}}$.

An example of choosing the prior model probabilities:

- $P(\gamma_0 = 0) = P(\mu = 0) = 3/4.$
- Independently, $P(\boldsymbol{\gamma}^* = \mathbf{0}) = 2/3$ and $\boldsymbol{\gamma}^* \neq \mathbf{0}$ have probability

  $$P(\boldsymbol{\gamma}^*) = \frac{26}{75} \cdot \frac{Beta(1+r, 1+25-r)}{Beta(1,1)}\,,$$

  where $r = \#\,\text{zeroes in } \boldsymbol{\gamma}^*$.
- Note that then
  - $P(\text{no effect}) = P(\mu = 0, \boldsymbol{\gamma}^* = \mathbf{0}) = 1/2$
  - $P(\mu \neq 0, \boldsymbol{\gamma}^* = \mathbf{0}) = 1/6$
  - $P(\mu = 0, \boldsymbol{\gamma}^* \neq \mathbf{0}) = 1/4$
  - $P(\mu \neq 0, \boldsymbol{\gamma}^* \neq \mathbf{0}) = 1/12$
  - $P(\gamma_i \neq 0) = 13/75$

The experimenter could (pre-experimentally) make different choices here to reflect beliefs as to which subgroups might most likely exhibit an effect, as long as $P(\text{no effect})$ is kept at 1/2. Post-experimentally, one cannot allow the experimenter to choose the prior probabilities of subgroups.

Possible Bayesian outputs of interest:

- $P(\text{effect of factor } i \neq 0 \mid \boldsymbol{x}) = \sum_{\{\boldsymbol{\gamma}:\gamma_i=1\}} P(\mathcal{M}_{\boldsymbol{\gamma}} \mid \boldsymbol{x})$.
- $P(\text{effect in subgroup } i \neq 0 \mid \boldsymbol{x}) = \sum_{\{\boldsymbol{\gamma}:\gamma_0=1 \text{ or } \gamma_i=1\}} P(\mathcal{M}_{\boldsymbol{\gamma}} \mid \boldsymbol{x})$.
- $P(\text{a constant effect} \neq 0 \mid \boldsymbol{x}) = P(\mathcal{M}_{(1,\mathbf{0})} \mid \boldsymbol{x})$.

Of course, posterior densities for all effects, conditional on their being nonzero, are also available.

Aggregation-based approaches

Basic idea: Recall that for every intersected subgroup (e.g., young, male, diabetic, non-smoker,...) there is an unknown mean $\mu_{\boldsymbol{v}}$. Plausible models involve aggregation of these means into common effects, e.g. $\mu_{\boldsymbol{v}_1} = \mu_{\boldsymbol{v}_2}$. There are a number of ways to aggregate means, including

- Product partition models (Hartigan and Berry)
- Dirichlet process models (Gopalan and Berry use for multiplicity control)
- Generalized partition models
- Species sampling models
- Tree-based models (our current favorite)

Surmountable problem: Any of these aggregate means could be zero; with some work, this can typically be handled by adding "zero" to the list.

Harder problem: Not all (not even most) aggregations are sensible (e.g., $\mu_{F_1G_1} = \mu_{F_2G_2} \neq \mu_{F_1G_2} = \mu_{F_2G_1}$ versus $\mu_{F_1G_1} = \mu_{F_2G_1} \neq \mu_{F_1G_2} = \mu_{F_2G_2}$).

Summary (about multiplicity)

- Developing methods for controlling for multiplicity is a dramatically increasing need in science.
- Approaching multiplicity control from the Bayesian perspective has the attractions that
  - there is a single approach that can be applied in any situation;
  - since multiplicity is controlled solely through prior probabilities of models, it does not depend on the error structure of the model;
  - there is flexibility in the assignment of model prior probabilities;
    - subjective assignments are pre-experimentally encouraged, to bring the science into the problem;
    - post-experimental objective assignments are also possible to evaluate "discovered" effects.
- Associated empirical Bayes analysis exhibits multiplicity control, but cannot be assumed to be an approximation to the Bayesian analysis.
- Bayesian subgroup analysis is promising, but challenging.

Thanks!