USING BAYESIAN CALCULATIONS TO ADDRESS CLINICAL TRIAL ISSUES DIRECTLY

A. Lawrence Gould
Merck Research Laboratories

NY Chapter, ASA Spring Symposium
June 16, 2003
Piscataway, NJ

goulda@merck.com

INTRODUCTION (1)

- Conventional clinical trial data analyses
  - confirmatory (hypothesis-testing)
  - use "tried-and-true" methods (ANOVA)
  - increasing acceptance of "exact" tests, generalized linear models, bootstrapping, etc.
- Hypothesis testing does not address all important issues
- Conventional methods often dictated by computing capability and need to use commercially available software
- Ultimately, only questions that conventional tools can answer get asked

INTRODUCTION (2)

- Many important questions involve estimation especially when trials include active controls
- Example:
  - Two trials compare treatments wrt preventing adverse event occurrence
  - Trial 1 rejects the effect equality, 95% confidence interval for the event rate ratio = (1.02, 1.07)
  - Trial 2 does not reject effect equality and 95% ci = (0.9, 10)
- Effect magnitude estimate with a measure of precision provides perspective on statistical and clinical significance

INTRODUCTION (3)

- Computational tools simplify implementing Bayesian methods
- Illustrate application of simple Bayesian methods using modern computational tools to answer real clinical questions
  - All applications to real clinical data
  - Two have been used as part of regulatory submissions – imputing or predicting a placebo response in an active-controlled trial
  - Proportion of treatment effect on the occurrence of a clinical event predictable by surrogate markers
- See also books by Berry & Stangl, Carlin & Louis, also BUGS manual

IMPUTATION (1)

- If active-controlled trial had included a placebo group, then reject H_0: active = placebo when statistic

$$t = t(\mathbf{x}_A, \mathbf{x}_P) > t_c$$

- If no placebo group, then t is random because the placebo group response is random

$$t = t(\mathbf{x}_A, \mathbf{X}_P)$$

How to find distribution of $\mathbf{X}_P$?

IMPUTATION (2)

- Suppose there are "similar" previous trials with observations on placebo

$f(\mathbf{x}_{Pi};\theta_i) \equiv$ likelihood for trial i, i = 1,…,q

- Assume θ_i varies among the previous trials with distribution $g(\theta;\Theta)$
- Marginal distribution for trial i is

$$h(\mathbf{x}_{Pi};\ \Theta) = \int f(\mathbf{x}_{Pi};\theta)\ g(\theta;\Theta)\ d\theta$$

- EMPIRICAL BAYES: Maximize product of marginal densities wrt elts of Θ, get maximum likelihood estimate $\hat{\Theta}$, assume mv normal in "large" samples
- The conditional predictive density of future values of $\mathbf{X}_P = h(\mathbf{x}_P;\ \hat{\Theta})$
- $\Pr(\text{Reject } H_0 |\ \hat{\Theta}) = \int_{t(x_A, x_P) > t_c} h(x_P;\hat{\Theta})dx_P$

IMPUTATION (3)

- STANDARD BAYES: As for EB, except add a prior for Θ, say $k_0(\Theta)$ with known parameters
- Posterior distribution of Θ:

$$k(\Theta;\text{data}) = h(x_{Pi};\Theta)k_0(\Theta) / \int \prod_i h(x_{Pi};\Theta)k_0(\Theta)d\Theta$$

- Given $k(\Theta;\text{data})$, the unconditional predictive distribution for $\mathbf{X}_P$ is

$$p(\mathbf{x}_P \mid \text{data}) =$$

$$\int_\Theta \int_\theta f(\mathbf{x}_P;\theta)\ g(\theta;\Theta)\ k(\Theta|\text{data})\ d\theta d\Theta$$

- Probability of rejecting H_0 then is

$$\int_{t(x_A, x_P) > t_c} p(\mathbf{x}_P \mid \text{data})\ d\mathbf{x}_P$$

Usually calculate this numerically using Markov Chain Monte Carlo (e.g., BUGS)

H_2-ANTAGONISTS IN DUODENAL ULCER (1)

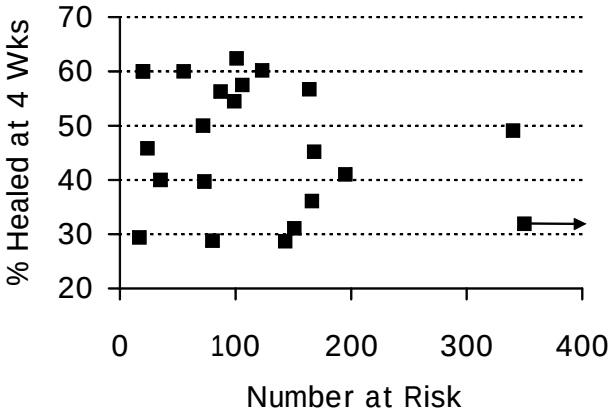
- Trial Outcome – Ulcer healing at 4 wks
- No placebo group, none of the between-group differences significant

		New H_2 Antagonist			Marketed
		Dose 1	Dose 2	Dose 3	Product
	At Risk	240	247	247	246
Healed	No.	164	191	201	186
Ulcer	%	68	77	81	76

- Question of interest: "How likely is it that findings reflect true efficacy as opposed to a high spontaneous healing rate?

H_2-ANTAGONISTS IN DUODENAL ULCER (2)

- Fortunately, lots of data on placebo from earlier trials, considerable variation in placebo response:



H_2-ANTAGONISTS IN DUODENAL ULCER (3)

- If trial had a placebo group, then could use simple chi-square (critical value = ξ) to test for active-placebo difference:

  Active: x of m healed

  Placebo: y of n healed

  $\Rightarrow$ Signif. active-placebo difference if $y \geq y_U$ or $y \leq y_L$, where y_L, y_U =

$$\frac{mn(2Nx-2\xi x+N\xi)\pm\sqrt{m^2n^2\xi^2+4mn\xi Nx(m-x)}}{2m(mN+n\xi)}$$

- Need to determine probability of a Y value falling outside (y_L, y_U), because trial did not include placebo

H_2-ANTAGONISTS IN DUODENAL ULCER (4)

- If Y for any trial is binomial with rate parameter p, and if values of p vary among trials with beta distn, then marginal distn of Y is beta-binomial:

$$p(y;n,a,b)=\binom{n}{y}B^{-1}(a,b)B(a+y,b+n-y)$$

- EMPIRICAL BAYES: Get ML estimates of a & b. Substitute these to get conditional predictive density of a future Y obsn given n and estimates
- For placebo data ML estimates are

$$\hat{a} = 9.28,\ \hat{b} = 11.2$$

H_2-ANTAGONISTS IN DUODENAL ULCER (5)

- Probabilities of rejecting H_0 can be calculated assuming n = 240 patients on placebo
- Maybe spuriously strong – conditional on values of a & b
- BAYES approach is unconditional; more conservative
- Comparative results for Empirical Bayes & Bayes analyses:

Prob(Reject H_0 | data)

	Dose 1	Dose 2	Dose 3	Mktd Prd.
Emp. Bayes	0.898	0.984	0.995	0.976
Bayes	0.887	0.974	0.988	0.964

H_2-ANTAGONISTS IN DUODENAL ULCER (6)

- Bayes: Assume logit $p_i = a + b_i$ (generalized linear model), $b_i \sim N(0,1/\tau)$

  Simple BUGS code

  2 runs, 1000 replicate burn-in for each, 8000 replicates thinned by ½, different starting points and random seeds

  Gave almost identical estimates
- Additional placebo healing rate variation reduces the confidence of significant active-placebo differences, but not by much
- Reasonable to assert that findings from this trial represent true activity rather than an unusually high placebo response rate.

ACE INHIBITORS IN CHF (1)

- Data from two trials of a new ACE Inhibitor (B = Bicycle, T = Treadmill)

Treatment	NEW ACEI	Placebo	NEW ACEI	'STD' ACEI
Study Control	Placebo	Placebo	Active	Active
Exercise Type	B or T	B or T	T	T
N	129	63	86	88
Mean	0.331	0.088	0.269	0.215
S.D.	0.381	0.307	0.277	0.297

- Essentially same protocol except for control and variations in exercise type
- Question of interest: What would result have been if active-control study had included a placebo group?

ACE INHIBITORS IN CHF (2)

- Placebo-controlled study data could have been used to answer this question, but doing so would not provide two independent "pivotal" studies.
- Fortunately, several studies in similar populations comparing ACEI vs placebo in literature
- Differed from protocols of current studies in some ways such as treatment duration and the exercise technique
- Account for exercise technique differences by using log of on-trt value/baseline value, assuming a modest change over time in the ability to perform work would give same % change for all techniques

ACE INHIBITORS IN CHF (3)

Exercise Tolerance Outcomes of Patients on Placebo from Prior Trials

Trial	Exercise Test Type	Trt Durn	No. Obs.	Mean	S.D.
1	Trdml, sec	12 w	42	0.048	0.392
2	Bike, watts	6 m	14	0.029	0.359
3	Trdml, sec	10 w	13	0.21	0.304
4	Bike, sec	3 m	8	-0.011	0.388
5	Bike, sec	3 m	8	-0.012	0.518
6	Bike, watts	6 m	8	0.016	0.347
7	Bike, sec	6 m	37	0.071	0.281
8	Bike, sec	12 w	11	0.098	0.502

- Reasonably consistent means and variances

ACE INHIBITORS IN CHF (4)

- Conclude active better than placebo if

$$t = (\bar{x}_A - \bar{x}_P) / \sqrt{s_A^2/n_A + s_P^2/n_P} > t_c$$

with $m = n_A + n_P - 2$ degrees of freedom

- No placebo group in active-controlled trial $\Rightarrow$ placebo group summary statistics unknown
- EMPIRICAL BAYES: Under normality, marginal density of sample mean & variance essentially product of chi-square & central t
  - use prior data to estimate parameters
  - integrate marginal density over rejection region to get [conditional] Pr(Reject H_0)

ACE INHIBITORS IN CHF (5)

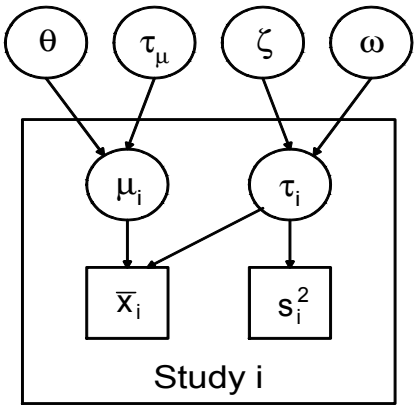
- Results of calculations

	Estimate	95% CI
μ_P	0.064	-0.28, 0.41
σ_P	0.36	0.25, 0.53
Pr(Reject H_0)	$ACEI_{New}$ 0.998	$ACEI_{Std}$ 0.981

- Variability of estimates $\Rightarrow$ est'd Pr(Reject H_0) may be optimistic

ACE INHIBITORS IN CHF (6)

- BAYES (unconditional) may be more realistic
- Graphical representation of model (also applies for Empirical Bayes)



ACE INHIBITORS IN CHF (7)

- Two runs done, with different starting values and different random seeds -- initial burn-in of 2000 reps followed by sets of 50,000 in which every 10^{th} result was retained
- Results:

	Estimate	95% CI
μ_P	0.054	-0.028, 0.13
σ_P	0.37	0.33, 0.42
Pr(Reject H_0)	$ACEI_{New}$ 0.922	$ACEI_{Std}$ 0.785

- Conclusion: New ACEI very likely would have differed significantly from placebo had a placebo group been included, probably std ACEI, too

Proportion of treatment effect explained by surrogate markers in HIV trial (1)

- Using surrogate markers as substitutes for clinical endpoints is risky

  Ideally should satisfy Prentice criteria but rare in practice
- Instead, consider how much markers explain treatment effect on clinical endpoints (PTE)
- PTE statistic for survival trials easy to calculate, but CI not – asymptotic result, software not readily available
- Possible to do Bayesian evaluation, get posterior dist'n, etc.

Proportion of treatment effect explained by surrogate markers in HIV trial (2)

- Use results from trial comparing a protease inhibitor alone and in combination with older drugs
  - frequent measurements of CD4+ cell counts and viral RNA levels
  - clinical endpoints = occurrence of AIDS-defining events
  - 964 patients in three treatment groups: PI monotherapy, PI combo therapy, and zudovidine (ZDV)
  - Weibull model describes clinical outcomes reasonably well (diagnostic)
  - Hazard function for Weibull model at time $t \propto \exp(z'\beta)$

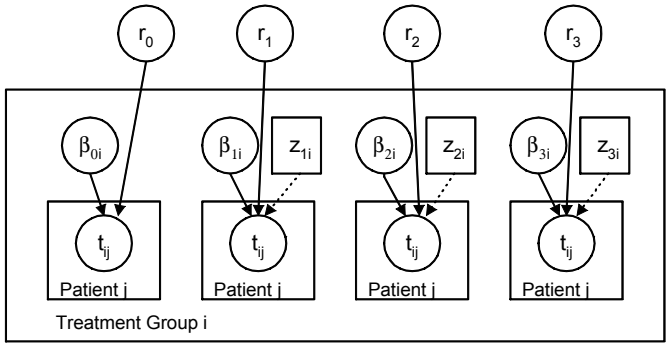
Proportion of treatment effect explained by surrogate markers in HIV trial (3)

- Values of PTE by cell count & viral load covariates ($CD4_0$, RNA_0, dCD4, dRNA):

	Cox	Weibull
Combo-ZDV	0.678	0.698
Mono-ZDV	0.759	0.793

Proportion of treatment effect explained by surrogate markers in HIV trial (4)

- Model used for Bayesian analysis of survival calculations



- Covariates used in models (plus trt)

Model	1	2	3	4
Covrs	none	$CD4_0$, dCD4	$CD4_0$, dCD4, dRNA	$CD4_0$, dCD4, RNA_0, dRNA

Proportion of treatment effect explained by surrogate markers in HIV trial (5)

- Focus on models 0 & 3

$$PTE = 1 - \left(\beta_{trt}^{(3)} - \beta_{ZDV}^{(3)}\right) \Big/ \left(\beta_{trt}^{(0)} - \beta_{ZDV}^{(0)}\right)$$

  trt = mono (PI) or combo (PI + ZDV)

- Calculate PTE for each set of trt effect realizations; these are draws from the PTE posterior distributions
- Results:

	Mean	Median	95% CI	Orig.
Mono – ZDV	0.769	0.789	0.188, 1	0.793
Comb - ZDV	0.647	0.668	0.088,1	0.698

Proportion of treatment effect explained by surrogate markers in HIV trial (6)

- PTE values consistent with conventional analysis results, but slightly lower
- Large variability $\Rightarrow$ processes may be occurring that the marker measurements do not capture
- PTE for PI alone > PTE for combo ?$\Rightarrow$ combo effect may be partly via pathways not mediated by CD4+ cell counts or viral RNA load

COMMENTS

- Computer-intensive Bayesian methods can be useful in pharmaceutical contexts, and use existing theory and software
- Many applications in other areas, including preclinical drug development
- Development of efficient methods for carrying out the calculations is a currently active area of research
- Evaluation of safety and tolerability in clinical trials – lack of significant trt-ctl differences usually interpreted as safety
- Absence of evidence is not evidence of absence – quantifying risk differences is more insightful and more relevant.