Why a Bayesian Be?

Andrew R. Willan, PhD
andy@andywillan.com

Senior Scientist
Program in Child Health Evaluative Sciences
SickKids Research Institute, Toronto

Professor of Biostatistics
Dalla Lana School of Public Health
University of Toronto

www.andywillan.com/talks

Three Main Reasons

Permits simple, intuitive and relevant statements of statistical inference
(Bayes Lite)

Provides a transparent framework for combining new information with current knowledge
(Bayes)

Facilitates decision theory for optimal decision-making and research design
(Full-on Bayes)

Simple, Intuitive, Relevant Statements of Statistical Inference

Frequentist definition of probability of an event: the limiting relative frequency of its occurrence in a series of repeated observations of a chance outcome in which it could occur.

For the Bayesian probability is the (subjective) expression of the uncertainty or "degree of belief" regarding the unknown.

Frequentist Statistical Inference

Frequentist definition of probability leads to the use of test of hypothesis, with associated *p*-values and confidence intervals, to characterize uncertainty regarding model parameters.

Working hypothesis

Null hypothesis (*i.e.* working hypothesis is not true)

If observations refute null hypothesis, the working hypothesis is "proven"

In empirical research, "refute" means observations are "unlikely" if null hypothesis is true. "Unlikely" usually means a probability less than 5%

Bayesian Statistical Inference

Bayesian definition of probability leads to the use of probability statements regarding model parameters to characterize the uncertainty.

Bayesian inference provides probability statements about the truth, given the data. Frequentist inference provides probability statements about the data, given the truth.

Consider a clinical trial comparing *T* and *S* with respect to the relative risk for a bad outcome, where the frequentist's *p*-value is 0.035 and where a one-sided test of hypothesis is applied at the 5% level

Simple, Intuitive, Relevant Statements of Statistical Inference

The frequentist statement of inference is:

"We can reject the null hypothesis that the relative risk is equal to or greater than one (*i.e. T* is equivalent or inferior to *S*) in favour of the alternative (working) hypothesis that the relative risk is less than one (*i.e. T* is superior to *S*) with a probability of being wrong is less than 5%.

This means that if the null hypothesis is true (*i.e. T* is equivalent or inferior to *S*) and the trial was repeated many, many times, the proportion of times that the results of these replications will be at least as inconsistent with the null hypothesis as the data from the trial under consideration is less than 5%."

Simple, Intuitive, Relevant Statements of Statistical Inference

This is not a statement about falsely rejecting the null hypothesis for this particular trial, but rather a statement about the proportion of many, many null hypothesis that would be falsely rejecting using the same criterion.

The Bayesian statement of inference is:

"The probability that the relative risk is less than one (*i.e. T* is superior to *S*) is 96.5%."

Simple, Intuitive, Relevant Statements of Statistical Inference

Frequentist 95% confidence interval:

"The 95% confidence interval for the relative risk is (0.493, 0.917), meaning that if the trial was conducted many, many times, then in the limit the proportion of the confidence intervals from these replications that include the true relative risk is 95%." The inference does not say anything about the probability that the confidence interval based on the data from this trial includes the true value of relative risk.

On the other hand, inference based on a Bayesian credible interval with the same limits is stated as, "there is a 95% probability that the relative risk lies in the interval (0.493, 0.917)."

Vitamin C & E for Preventing Adenomas

Greenberg *et al*. A Clinical Trial of Antioxidant Vitamins to Prevent Colorectal Adenoma. *NEJM* 1994; **331**(4):141-147.

Factorial, placebo controlled RCT of 751 patients who had had a previous adenoma removed.

Essentially four arms:

1. Placeboes
2. Beta Carotene
3. Vitamins C & E
4. Beta Carotene and Vitamins C & E

Vitamin C & E for Preventing Adenomas

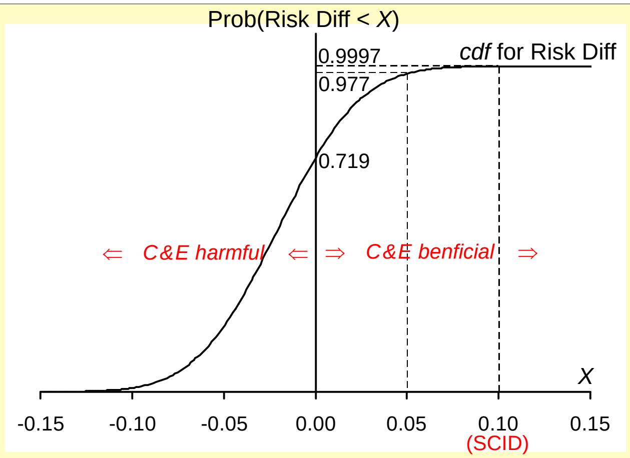
Combining arms 1 & 2 versus 3 & 4 to examine effect of vitamins C & E

Arms 1 & 2 (no vitamins): 134 of 371 (36.1%) had adenomas
Arms 3 & 4 (vitamins C & E): 145 of 380 (38.2%) had adenomas

Estimate of Risk Difference (risk on Arms 1 & 2 *minus* risk on Arms 3 & 4):
0.361 – 0.382 = -0.021 (2p-value: 0.563)

95% Confidence Interval: -0.09 to 0.048

Vitamin C & E for Preventing Adenomas



Probability that Risk Difference is less than SCID is 0.9997

Steroids for Women at Risk of Early Delivery

Numerous RCTs have found that a single course of steroids for pregnant women who are at risk of early delivery reduces the risk of respiratory distress syndrome and other bad outcomes in the babies.

Only three have reported on long-term neurological abnormalities.

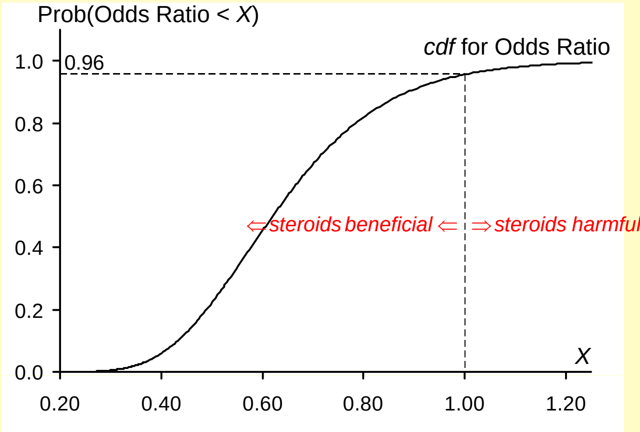
This is important since it relates to a lawsuit.

Steroids for Women at Risk of Early Delivery

A meta-analysis of long-term neurological abnormalities reports an odd ratio of 0.62, with a 95% confidence interval of 0.36 to 1.08 and a two-sided *p*-value of 0.08.

Led people to make statements such as: "no evidence of long-term benefit."

Steroids for Women at Risk of Early Delivery



Probability that steroids reduces the risk of long-term neuro. ab. is 0.96.

Just Published (no-name) Example

T: *n* = 765; bad outcome = 398 (52.0%)
S: *n* = 767; bad outcome = 430 (56.1%)

Two-sided *p*-value: 2*p* = 0.12

Meta-analysis with previous pilot yields 2*p* = 0.06

Conclusion: *T* "does not reduce the rate of" the bad outcome.

Bayesian Conclusion: The probability that *T* reduces the probability of a bad outcome is 97%.

Framework for Combining New Information (Data) with Current Knowledge

Bayes Theorem: $p(\theta \mid y) = \frac{p(y \mid \theta)}{p(y)} p(\theta)$

$p(\theta)$ expresses the current knowledge as a probability distribution function (*pdf*) for the model parameters, referred to as the prior distribution

$p(y \mid \theta)$ is the likelihood for the data, *i.e.* the *pdf* for the data, given θ

$p(\theta \mid y)$ is the posterior *pdf* for θ given the data

$p(y)$ is a normalizing factor to ensure that $p(\theta \mid y)$ integrates to 1

Facilitates Decision Theory

In the face of uncertainty, decision theory permits optimal decision making, answering questions, such as:

Should a new intervention be adopted for future patients?

Is more research needed?

If so, how big should the study be?

Facilitates Decision Theory

A new intervention should be adopted if no more research is needed

More research needed if the value of the information from the research is greater than its cost

The size of the study should maximize the difference between the value and the cost

What is the Value of Information?

Faced with making a decision based on current information and the associated uncertainty, what is the expected value of additional information?

It is the amount by which the new information reduces the expected opportunity loss of making a particular decision

Okay, So What is the Opportunity Loss and How Does One find Its Expected Value?

The opportunity loss is the utility of the best decision *minus* the utility of the decision made

The opportunity lost is a function of the “truth” and one finds its expected value by using the prior *pdf* of the “truth” given by the current information

Expected opportunity loss (EOL) is A.K.A. the expected value of perfect information (EVPI), because if you had perfect information, you could avoid the opportunity loss

That all very well, but can we have an example please?

Suppose you are told the one of four boxes contains $1000; the rest are empty. You get to choose one box keep its contents.

Suppose further that you have the opportunity to pay $200 to have revealed to you two of the <u>empty</u> boxes.

Is the expected value of this information (EVSI) worth $200?

i.e. Is it worth paying $200 to increase the probability of choosing the right box from 0.25 to 0.5?

Example Continued

Opportunity loss = Utility of best choice – Utility of choice made
= $1000 – contents of box chosen

Box	Contents	Opp. Loss	Prior Info.	Prior Exp. Opp. Loss	Post Info.	Post Exp. Opp. Loss
A	0	1000	0.25	250	0	
B	1000	0	0.25	0	0.5	0
C	0	1000	0.25	250	0.5	500
D	0	1000	0.25	250	0	
				750		500

$\text{EVSI} = \text{EOL}_{prior} - \text{EOL}_{post} = 750 - 500 = 250$

Prostate Cancer Trial

Mitoxantrone + Prednisone (*T*) versus Prednisone alone (*S*) for symptomatic hormone resistant prostate cancer

161 patients

No difference in survival

Better palliation with *T*

Cost data on 114 patients from the 3 largest centres

Retrospective chart review; included hospital admissions, outpatient visits, investigations, therapies and palliative care

Quality-adjusted survival using EORTC QLQ-C30

Prostate Cancer Trial

Let e_j and c_j be the mean quality-adjusted survival time and cost for arm j (=T,S)

Then $\Delta_e = e_T - e_S$ and $\Delta_c = c_T - c_S$

Incremental Net Benefit (INB): $b(\lambda) = \Delta_e \lambda - \Delta_c$, where λ is the threshold value for a unit of effectiveness

$$b(\lambda) = (e_T - e_S)\lambda - (c_T - c_S) = e_T\lambda - c_T - (e_S\lambda - c_S) = \text{NB}_T - \text{NB}_S$$
$$\therefore \text{NB}_S - \text{NB}_T = -b(\lambda)$$

Prostate Cancer Trial

	Treatment	Standard		
$\bar{e}_j$	40.9	28.1	difference = $\hat{\Delta}_e$ = 12.8	(QALW)
$\bar{c}_j$	27,322	29,039	difference = $\hat{\Delta}_c$ = -1717	(CDN $)
$\hat{V}(\bar{e}_j)$	24.1	16.4	sum = $\hat{V}(\hat{\Delta}_e)$ = 40.5	
$\hat{V}(\bar{c}_j)$	6,466,351	7,872,681	sum = $\hat{V}(\hat{\Delta}_c)$ = 14,339,032	
$\hat{C}(\bar{e}_j,\bar{c}_j)$	2771	2876	sum = $\hat{C}(\hat{\Delta}_e,\hat{\Delta}_c)$ = 5647	

$$\hat{b}(\lambda) = \hat{\Delta}_e\lambda - \hat{\Delta}_c = 12.8\lambda + 1717$$

$$\hat{V}\left(\hat{b}(\lambda)\right) = \hat{V}(\hat{\Delta}_e)\lambda^2 + \hat{V}(\hat{\Delta}_c) - 2\hat{C}(\hat{\Delta}_e,\hat{\Delta}_c)\lambda = 40.5\lambda^2 + 14339032 - 11294\lambda$$

Prostate Cancer Trial

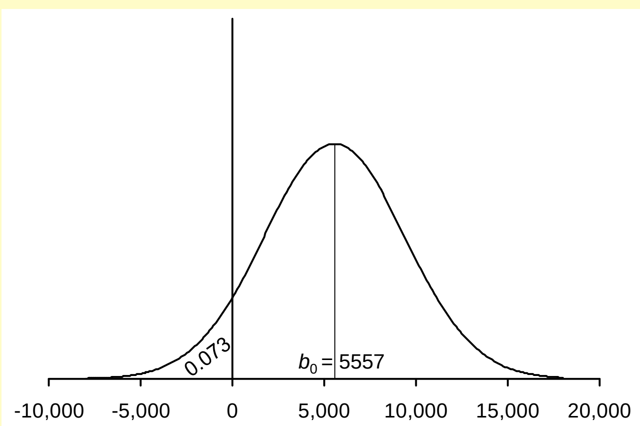
Assuming no prior information, the posterior distribution for INB is

$$b(\lambda) \sim N\left(12.8\lambda + 1717;\, 40.5\lambda^2 + 14339032 - 11294\lambda\right)$$

For $\lambda = 300$,

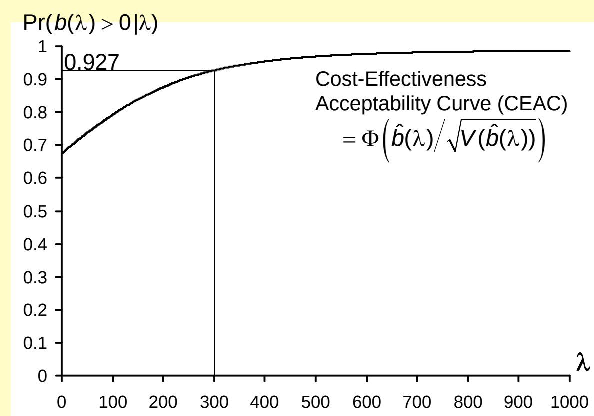
$$b(300) \sim N(5557;\, 14{,}595{,}832) = N(b_0; v_0)$$

Prostate Cancer Trial



Posterior *pdf* for incremental net benefit: $b(300)$

Prostate Cancer Trial



$\Pr(b(\lambda) > 0|\lambda=300) = 0.927$

significance for (H: $b(300) \le 0$ vs. A: $b(300) > 0$) = 1 – 0.927 = 0.073

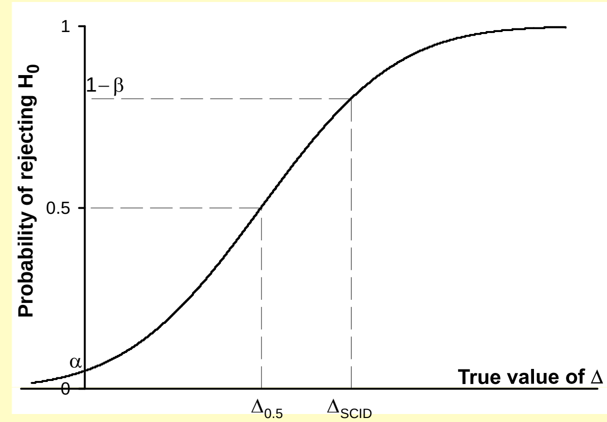
What to do for λ = 300?

Classical statistician: "Sorry folks, p = 0.073 > 0.05, stick with *Standard*.
Next project, please."

Bayesian statistician: "Whadayanuts!?
The probability that *Treatment* is cost-effective is 92.7%.
It's criminal to stick with *Standard*, adopt *Treatment*."

Informed (*i.e.* decision theory aware) clinical trial methodologist:
"Perhaps we need more information (*i.e.* another trial).
The size of trial (*n*/arm) should maximize the difference between the value of doing the trial and the cost of doing it."

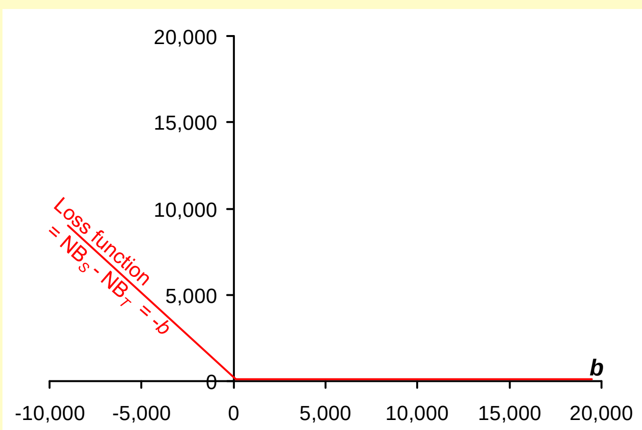
Traditional Power Curve and Sample Size Determination



H_0: $\Delta \le 0$, *i.e. Treatment* equal or inferior to *Standard*

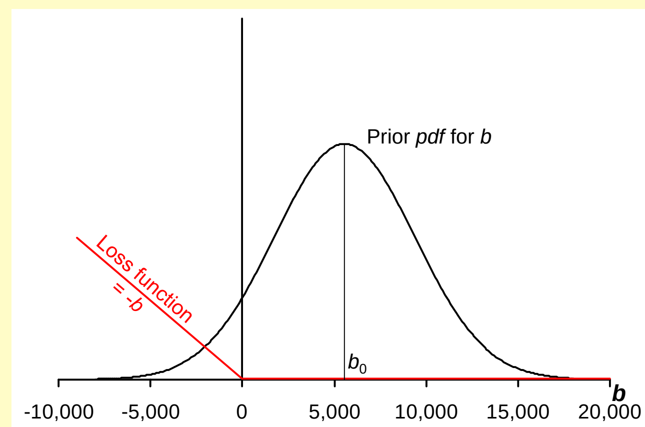
Rejection of H_0 is argument for adopting *Treatment*

Prostate Cancer Trial



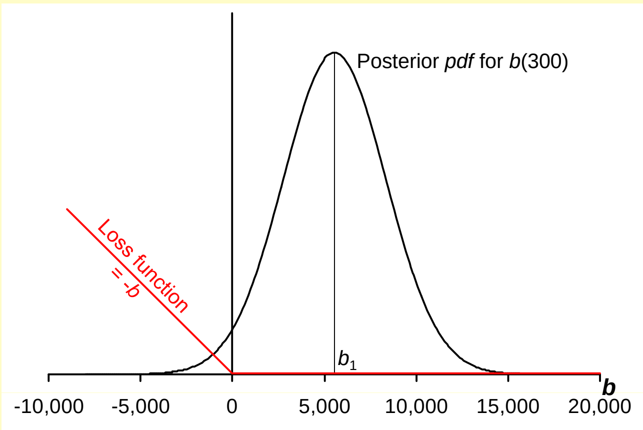
Loss function for adopting *T* = net benefit(best) – net benefit(*T*)

Prostate Cancer Trial



Pre-trial EOL/patient = function of b_0, v_0

Prostate Cancer Trial



Post-trial EOL/patient = function of b_0, v_0, n, σ_+^2

Prostate Cancer Trial

EVSI/patient = EOL/patient$_{pre}$ – EOL/patient$_{post}$ > 0

Some assumptions

- Time horizon (*h*) = 20 years
- Incidence (*k*) = 2500/year
- Accrual rate (*a*) = 2500/year
- Results are instantaneous (τ = 0)
- Perfect implementation, *i.e.* if INB > 0 and no further evidence is sought or expected all patients would receive *Treatment*
- No discounting
- No cost of adoption
- Equal sample size per arm
- No interim analysis
- No between trial variation in mean INB

Prostate Cancer Trial

Expected Value of Sample Information for the trial:

EVSI(n) = $(hk - 2n)$EVSI/patient

Expected Total Cost for the trial: ETC(n) = $C_f + 2nC_v + nb_0$

where

- n is the number of patients on each arm
- C_f is the fixed cost
- C_v is the variable cost per patient

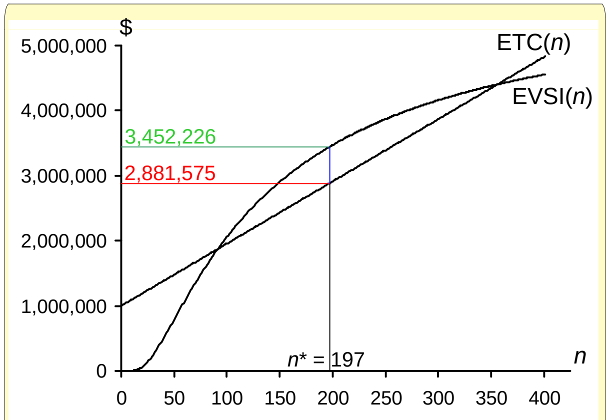
Expected Net Gain: ENG(n) = EVSI(n) – ETC(n)

Let n^* maximize ENG(n)

If ENG(n^*) > 0, then optimal sample size is n^*

If ENG(n^*) <= 0, then optimal sample size is 0

Prostate Cancer Trial



Time horizon: h = 20 yrs

Incidence: k = 2500/yr

Accrual: a = 2500/yr

Follow-up/Analysis τ = 0

C_f = 1,000,000

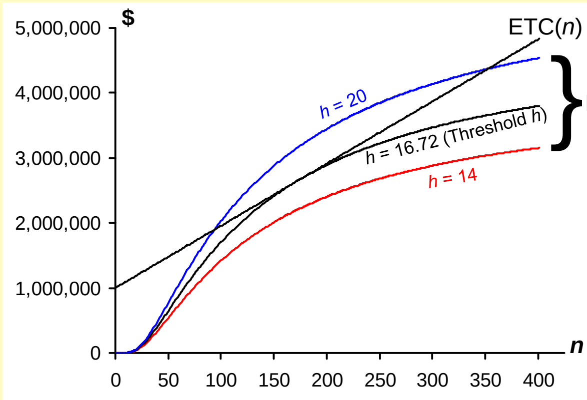
C_v = 2000

Financial cost = 1,788,000; E(opportunity cost) = 1,093,575;

ETC(197) = 2,881,575; EVSI(197) = 3,452,226

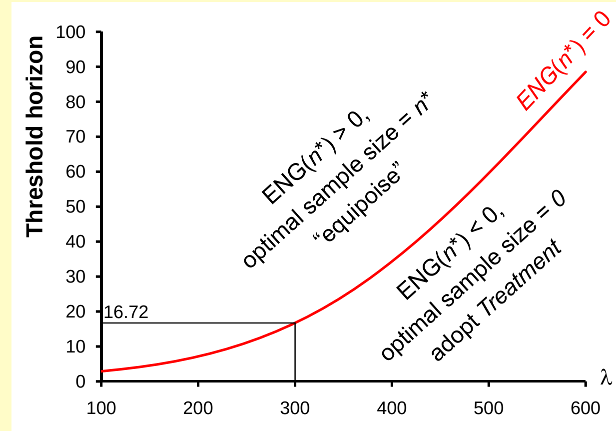
ENG(196) = 570,651

Prostate Cancer Trial



EVSI(n) = (hk – $2n$) x EVSI/patient

Prostate Cancer Trial



Combinations horizon (h) and λ for which ENG(n^*) = 0

Summary—Why A Bayesian Be?

Permits simple, intuitive and relevant statements of statistical inference

Provides a transparent framework for combining new information with current knowledge

Facilitates decision theory for optimal decision-making and research design

References—Why A Bayesian Be?

- Spiegelhalter DJ, Abrams KR, Myles JP .2004. *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. Chichester UK: Wiley.
- Ades AE, Lu G, Claxton K. 2004. *Medical Decision Making* **24**: 207-227.
- Willan AR, Pinto EM. 2005. *Statistics in Medicine* **24**: 1791-1806.
- Eckermann S, Willan AR. 2007. *Health Economics* **16**: 195-209.
- Willan AR. 2007. *Clinical Trials* **4**: 279-285.
- Willan AR. 2008. *Clinical Trials* **5**: 587-594.
- Willan AR, Kowgier ME. 2008. *Clinical Trials* **5**: 289-300.
- Eckermann S, Willan AR. 2008. V*alue in Health* **11**: 522-526.
- Eckermann S, Willan AR. 2008. *Medical Decision Making* **28**: 300-305.
- Eckermann S, Willan AR. 2009. *Health Economics* **18**: 203-216.
- Willan AR, Eckermann S. 2010. *Health Economics* **19**:549-561
- Willan AR, Eckermann S. 2011. Expected value of information and pricing new health care interventions. *PharmacoEconomics* (accepted).
- Willan AR, Eckermann S. 20??. Accounting for between-study variation in value of information methodology. *Health Economics* (under review).