

Bayes meets the NEJM

(Or why the sun hasn't exploded)

Jay Brophy MEng MD FRCP FACC FCCS FCAHS PhD
Sept 21 2021

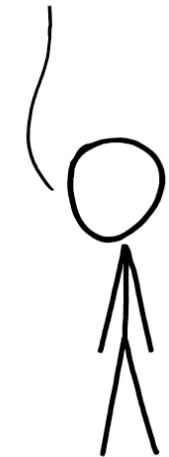
(IT'S NIGHT, SO WE'RE NOT SURE.)

THIS NEUTRINO DETECTOR MEASURES WHETHER THE SUN HAS GONE NOVA.

THEN, IT ROLLS TWO DICE. IF THEY BOTH COME UP SIX, IT LIES TO US. OTHERWISE, IT TELLS THE TRUTH.

LET'S TRY.

DETECTOR! HAS THE SUN GONE NOVA?



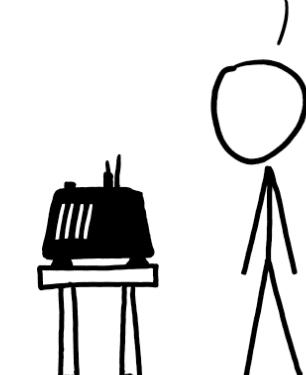
FREQUENTIST STATISTICIAN:

THE PROBABILITY OF THIS RESULT HAPPENING BY CHANCE IS $\frac{1}{36}=0.027$. SINCE $p<0.05$, I CONCLUDE THAT THE SUN HAS EXPLODED.



BAYESIAN STATISTICIAN:

BET YOU \$50 IT HASN'T.



Conflicts of Interest

I have **no known conflicts** associated with this presentation and to the best of my knowledge, am **equally disliked** by

- 1) all pharmaceutical and device companies
- 2) quite probably all NEJM editors (if they were made aware of this presentation)



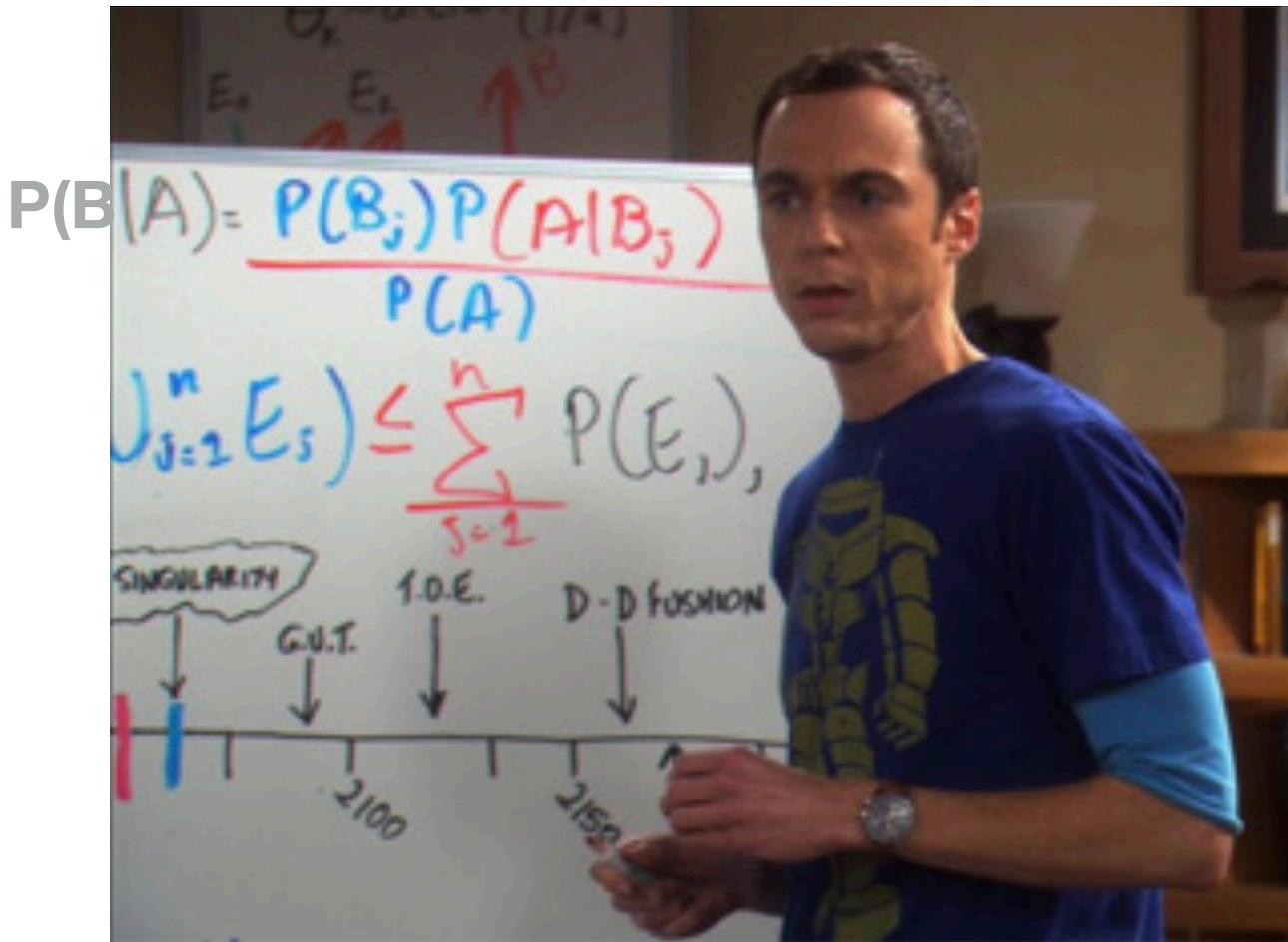
<http://www.nofreelunch.org/>

Objectives

1. Appreciate the limitations of standard statistical inference
2. Understand the general philosophy and basic mechanism of Bayesian inference
3. Appreciate the advantages of Bayesian inference in analyzing contemporary CV randomized clinical trials

Quiz #1

- Name the TV show and the main character
- What is “cool” about the image?



Quiz #2

- Consider two claims.

(1) Todd claims that he can predict coin flips. To test his claim, you flip a fair coin 10 times and he correctly predicts all 10.

(2) Emily claims that she can distinguish between natural and artificial sweeteners. To test her claim, you give her 10 sweetener samples and she correctly identifies all 10.

Given these experiments, which of the 2 statements below do you **most** agree with?

- A. The evidence supporting Todd's claim is just as strong as the evidence supporting Emily's claim.
- B. You are more confident in Emily's claim than Todd's claim.

Statistical Inference

Statistical Inference

- Statistical inference is the process of drawing conclusions or scientific truths (including understanding mechanisms that may have caused or generated that data) about populations from a data sample (which is usually noisy)
- Dependent on study design and **analysis**

Two views of probability

- Long run (Frequentist)
 - probability = long run frequency an event occurs in independent identically distributed (iid) repetitions
 - IOW, if an experiment is repeated many many times, the resultant percentage defines the **fixed**, but unknown, population parameter
- Subjective (Bayesian)
 - probability = subjective belief or relative plausibility based on prior beliefs and updated as additional evidence is acquired
 - IOW “probability relates partly to our ignorance, partly to our knowledge” - Laplace 1812
 - parameters are viewed as **random** variables with their own distributions

Frequentist Inference

- **Advantages**

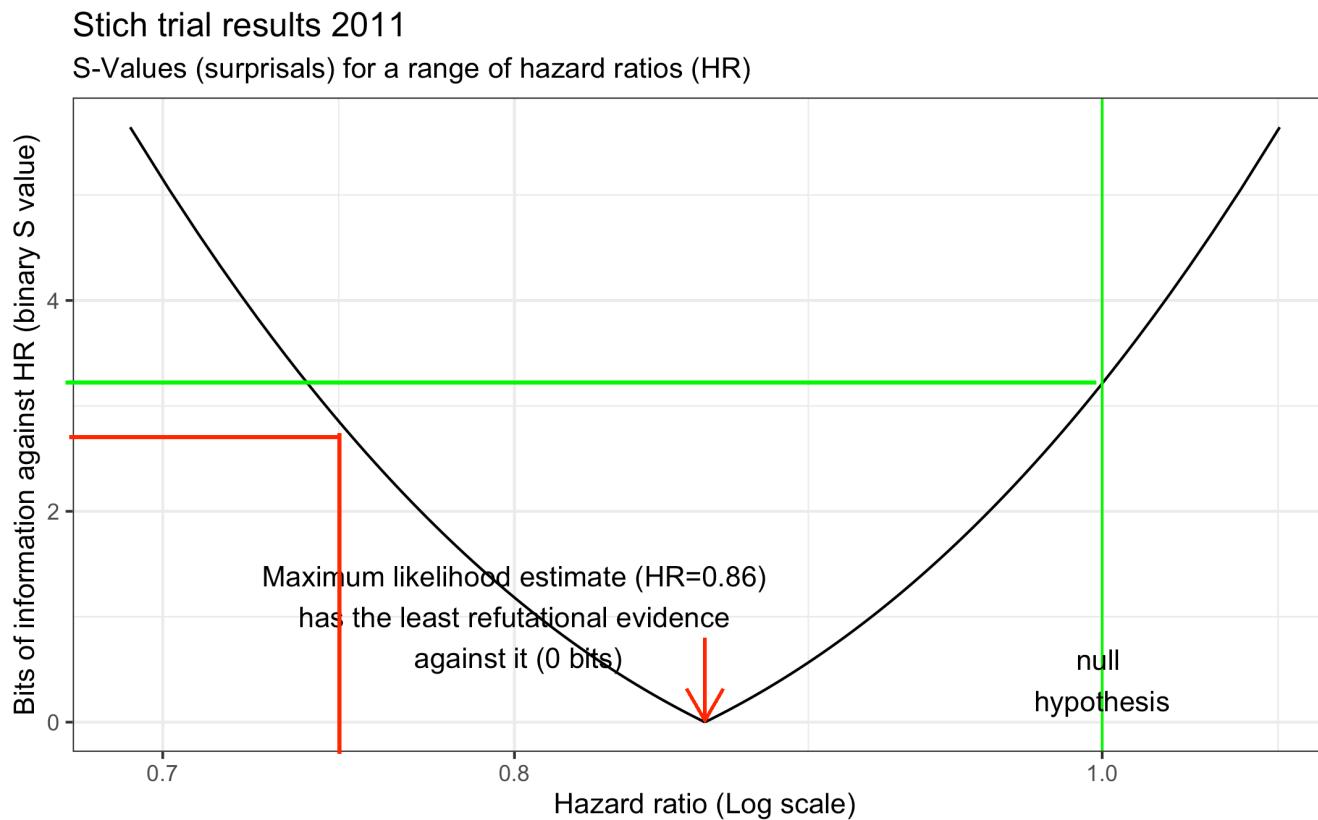
- most commonly taught, practiced, and accepted approach

- **Disadvantages**

- involves unholy marriage of different and conflicting statistical paradigms (Neyman Pearson (NHST) & Fisher (p values))
- prone to multiple misinterpretations
 - p value \neq probability H_0 true
 - $p = \text{NS}$ (absence of evidence \neq evidence of absence)
 - p value poor measure (over estimates) of strength of evidence
 - p value same from small effect and large study or with large effect in smaller study
 - often reduces scientific inference to a threshold decision ($p < 0.05$)
 - allocates excessive importance to the null hypothesis H_0

The problem of nullism

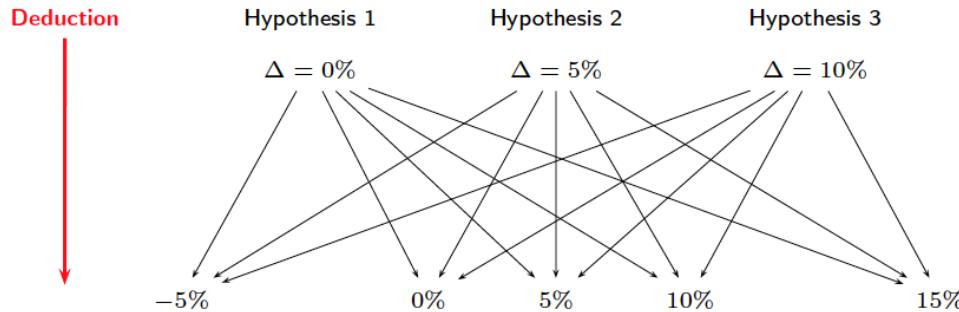
- **S value** ($\log_2(1/P)$) measures **strength of the evidence against** not only H_0 but against any specific H_a - binary scale (# heads coin tosses)
- Consider the result $HR = 0.86$ (95% CI 0.72 - 1.04, $p = 0.12$)



- < evidence against H_a = 25% decrease (red) than there is against H_0 hypothesis (green), which we have been told we should accept!

Deductive inference

Frequentist



Fix the hypotheses

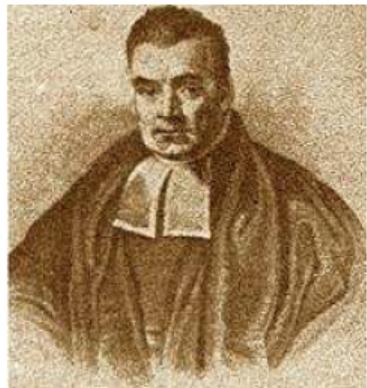
By deduction, make inference on the observed data

$\Pr(\text{Observed data} \mid \text{Hypothesis})$

If small, then deduce weak support of the evidence to the hypothesis

Bayesian Inference

Bayesian inference



Reverend Thomas Bayes (1702 - 1761)

P R O B L E M.

Given the number of times in which an unknown event has happened and failed: *Required* the chance that the probability of its happening in a single trial lies somewhere between any two degrees of probability that can be named.

In modern language: given $y \sim \text{Binomial}(\Theta, n)$,

find

$$\Pr(\Theta_1 \leq \Theta \leq \Theta_2 | y, n)$$

- **Bayes' Theorem** -> probability statements about hypotheses, model parameters or anything else that has associated uncertainty
- **Advantages**
 - allows consideration of complex questions / models where all sources of uncertainty can be simultaneously and coherently considered
 - provides direct and meaningful answers to research questions
 - allows integration of all available information
 - mirrors human learning with constant updating
- **Disadvantages**
 - subjectivity (?)
 - problem of induction (Hume / Popper - difficulty generalizing about future)

Acknowledging subjectivity (judgement)

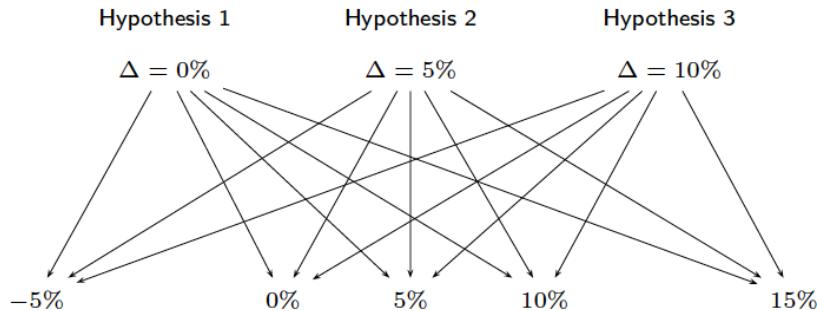
- Bayesian
 - in choice of priors
- Frequentist more examples of subjectivity, choice of
 - type 1 and type 2 error rates
 - 1 or 2 sided tests
 - stopping rules
 - model and covariates
 - multiplicity adjustments...
- Life
 - choice of partner, life insurance, career, music, art, books,...
- Subjectively is unavoidable, no need to be afraid of it
 - be as “objective”, “scientific” and “transparent” as possible
 - random guesses, data selectivity and biases to be avoided



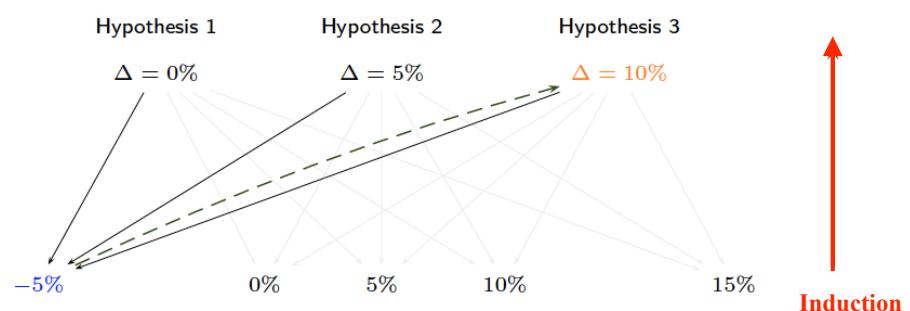
Deductive vs. inductive inference

Frequentist

Deduction



Bayesian



Fix the hypotheses

By deduction, make inference on the observed data

Pr(Observed data | Hypothesis)

If small, then deduce weak support of the evidence to the hypothesis

Fix observed data

By induction, make inference on unobservable hypotheses

Pr(Hypothesis | Observed data)

If less than the probability of other competing hypotheses, then weak support of the evidence to the hypothesis

Versions of Bayes' Theorem

- Conceptual



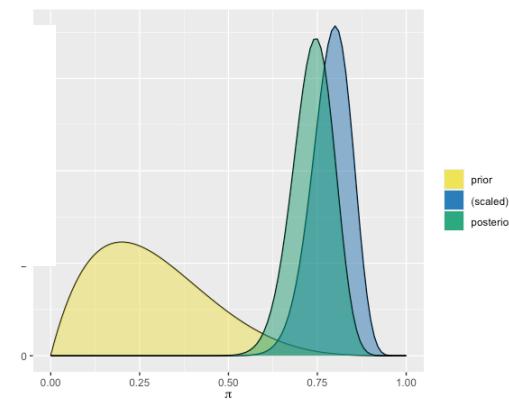
$$\text{posterior} = \frac{\text{prior} \cdot \text{likelihood}}{\text{normalizing constant}}$$

- Discrete

$$P(A | B) = \frac{P(B | A)P(A)}{P(B | A)P(A) + P(B | \bar{A})P(\bar{A})}$$

- Continuous

$$P(A|B) = \frac{P(B|A)P(A)}{\int_0^\infty P(B|A)P(A)dA}$$



Bayesian Inference & Terminology

- Combines **belief** (prior) & **evidential** (likelihood ratio) calculus

$$\text{posterior} = \frac{\text{prior} \cdot \text{likelihood}}{\text{normalizing constant}} \propto \text{prior} \cdot \text{likelihood}$$

Likelihood - propensity for observing a certain value of B given a certain value of A

Prior - what we know of A **before** seeing B

$$P(A|B) = \frac{P(B|A)P(A)}{\int_0^\infty P(B|A)P(A)dA}$$

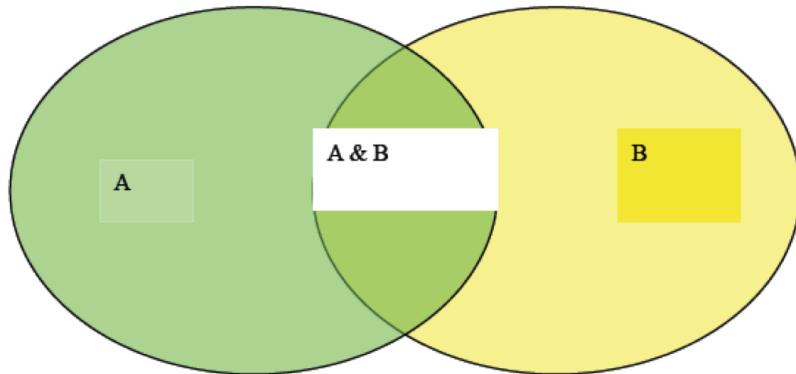
Posterior - what we know of A **after** seeing B

Normalizing constant - to ensure that the left hand side is a valid probability distribution

Bayes' Theorem follows basic probability rules

Multiplication rule

$$p(A|B) = p(A \cap B) / p(B)$$



Addition (total probability) rule

$$p(B) = p(B \cap A) + p(B \cap \bar{A})$$

$$p(B) = p(B | A) p(A) + p(B | \bar{A}) p(\bar{A})$$

- If A & B **independent** then $P(A | B) = P(A)P(B) / P(B) = P(A)$
 - i.e. knowing something about B, tells you nothing about A

Bayes' Theorem follows basic probability rules

Multiplication rule X 2

$$p(A|B) = p(A \text{ and } B) / p(B)$$

$$p(B|A) = p(A \text{ and } B) / p(A)$$

$$p(A|B) * p(B) = p(B|A) * p(A)$$

$$p(A|B) = \frac{p(B|A) * p(A)}{p(B)}$$

Addition rule

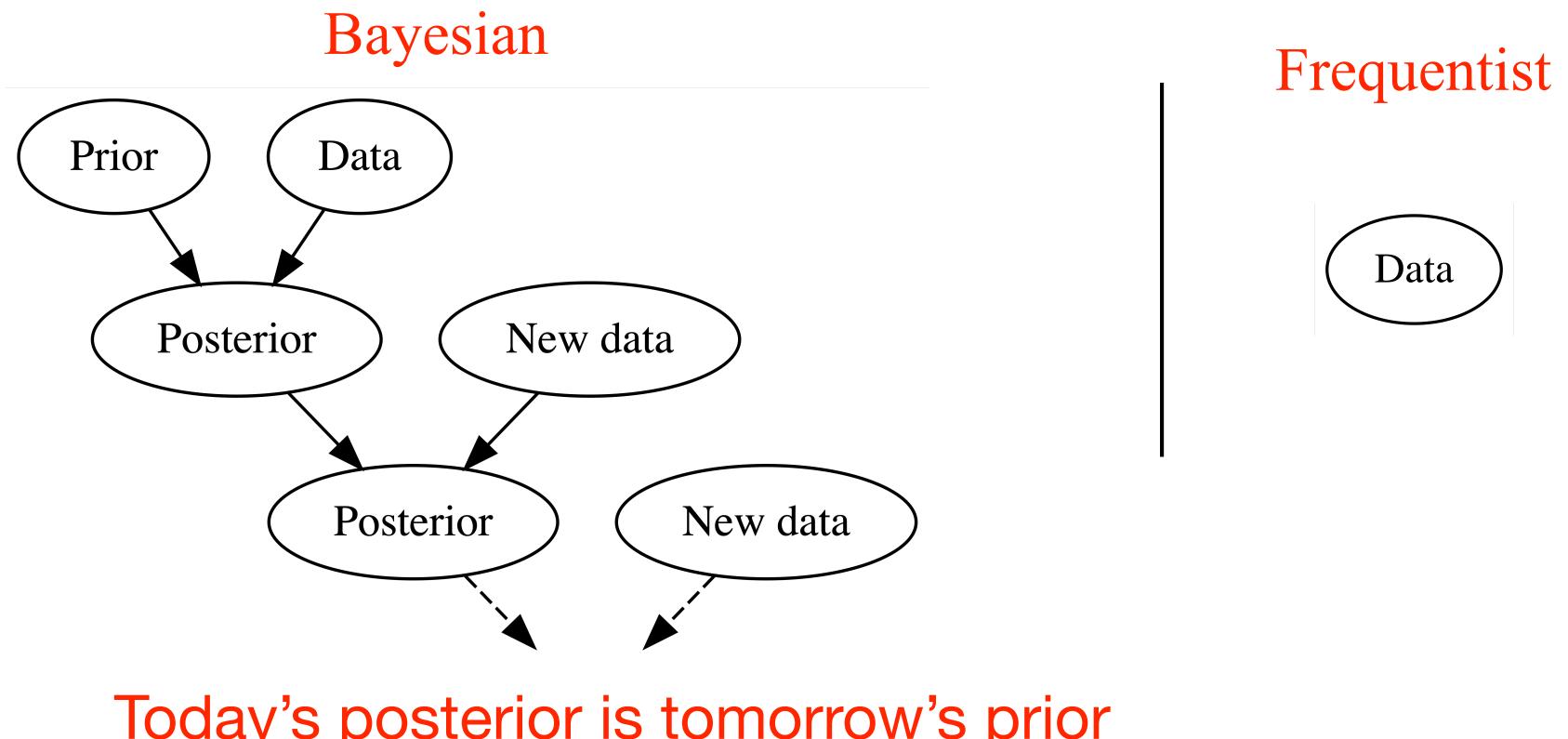
$$p(B) = p(B | A) p(A) + p(B | \bar{A}) p(\bar{A})$$

Bayes' Theorem
(discrete version)

$$P(A | B) = \frac{P(B | A)P(A)}{P(B | A)P(A) + P(B | \bar{A})P(\bar{A})}$$

Bayesian Inference

- Everyone is entitled to their own, subjective evaluation which can be updated with new evidence (data).
- In **identifiable** problems as more data accumulates the subjective component diminishes and divergent opinions converge
- In **non-identifiable** problems (missing data, measurement error, unmeasured confounders) priors can remain important even as more data accumulates



Bayes' Theorem as Odds Ratio

Posterior odds = likelihood ratio * Prior odds

Posterior probability = $\frac{\text{Posterior odds}}{1 + \text{Posterior odds}}$

Bayes' Theorem X 2

$$p(A|B) = \frac{p(B|A) * p(A)}{p(B)} \quad (\text{eq 1})$$

$$p(\bar{A}|B) = \frac{p(\bar{B}|\bar{A}) * p(\bar{A})}{p(B)} \quad (\text{eq 2})$$

Odds ratio

(eq 1 / eq 2)

$$\text{Posterior odds} = \frac{p(A|B)}{p(\bar{A}|B)} = \text{Likelihood Ratio} = \frac{p(B|A)}{p(B|\bar{A})} \quad * \quad \text{Prior odds} = \frac{p(A)}{p(\bar{A})}$$

Bayes factors

[BayesFactor user's manual](#)[BayesFactor on CRAN](#)[BayesFactor GitHub development](#)[BayesFactor on Facebook](#)

The **Bayes factor** is the relative evidence in the data. The evidence in the data favors one hypothesis, relative to another, exactly to the degree that the hypothesis predicts the observed data better than the other.

$$\frac{P(H_a | y)}{P(H_n | y)} = \frac{P(y | H_a)}{P(y | H_n)} \times \frac{P(H_a)}{P(H_n)}$$

Posterior odds = **BayesFactor** * **Prior odds**

BAYES FACTOR BF_{10}	LABEL
> 100	Extreme evidence for H1
30 – 100	Very strong evidence for H1
10 – 30	Strong evidence for H1
3 – 10	Moderate evidence for H1
1 – 3	Anecdotal evidence for H1
1	No evidence
1/3 – 1	Anecdotal evidence for H0
1/3 – 1/10	Moderate evidence for H0
1/10 – 1/30	Strong evidence for H0
1/30 – 1/100	Very strong evidence for H0
< 1/100	Extreme evidence for H0

P values & Bayes factors - exchange rate

$$\text{Minimum Bayes factor} = e^{-Z^2/2}$$

P Value (Z Score)	Minimum Bayes Factor	Decrease in Probability of the Null Hypothesis, %		Strength of Evidence
		From	To No Less Than	
0.10 (1.64)	0.26 (1/3.8)	75	44	Weak
		50	21	
		17	5	
0.05 (1.96)	0.15 (1/6.8)	75	31	Moderate
		50	13	
		26	5	
0.03 (2.17)	0.095 (1/11)	75	22	Moderate
		50	9	
		33	5	
0.01 (2.58)	0.036 (1/28)	75	10	Moderate to strong
		50	3.5	
		60	5	
0.001 (3.28)	0.005 (1/216)	75	1	Strong to very strong
		50	0.5	
		92	5	

Minimum BF (or minimum likelihood ratio) is the smallest amount of evidence that can be claimed for the null hypothesis (or the strongest evidence against it) on the basis of the data.

Ex. Prior odds 1, or prior probability = 50%
 $p = 0.05, BF = .15$
Post odds = .prior odds * BF = 1 * .15 = .15
Prob = Odds / (Odds +1) = .15 / 1.15 = 13%

The weight of evidence against the null hypothesis is not nearly as strong as the magnitude of the P value suggests & depends on your prior belief

Examples

Bayesian reanalysis of RCTs

Placing Trials in Context Using Bayesian Analysis

GUSTO Revisited by Reverend Bayes

James M. Brophy, MD, Lawrence Joseph, PhD

Standard statistical analyses of randomized clinical trials fail to provide a direct assessment of which treatment is superior or the probability of a clinically meaningful difference. A Bayesian analysis permits the calculation of the probability that a treatment is superior based on the observed data and prior beliefs. The subjectivity of prior beliefs in the Bayesian approach is not a liability, but rather explicitly allows different opinions to be formally expressed and evaluated. The usefulness of this approach is demonstrated using the results of the recent GUSTO study of various thrombolytic strategies in acute myocardial infarction. This analysis suggests that the clinical superiority of tissue-type plasminogen activator over streptokinase remains uncertain.

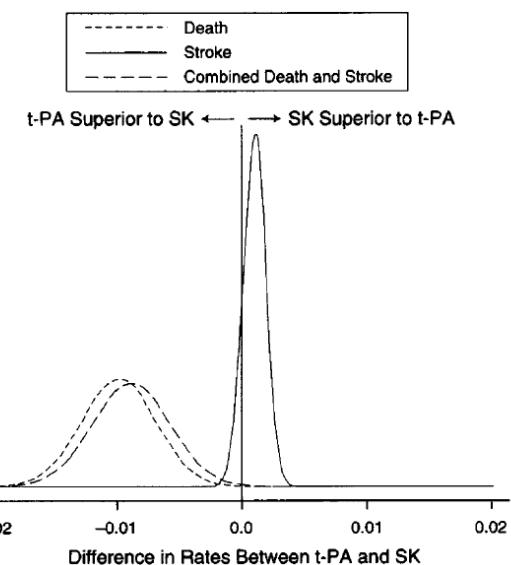
(JAMA. 1995;273:871-875)

Table 2.—Probability of t-PA Superiority as a Function of Prior Belief in GISSI-2 and ISIS-3 Data After Consideration of the GUSTO Data*

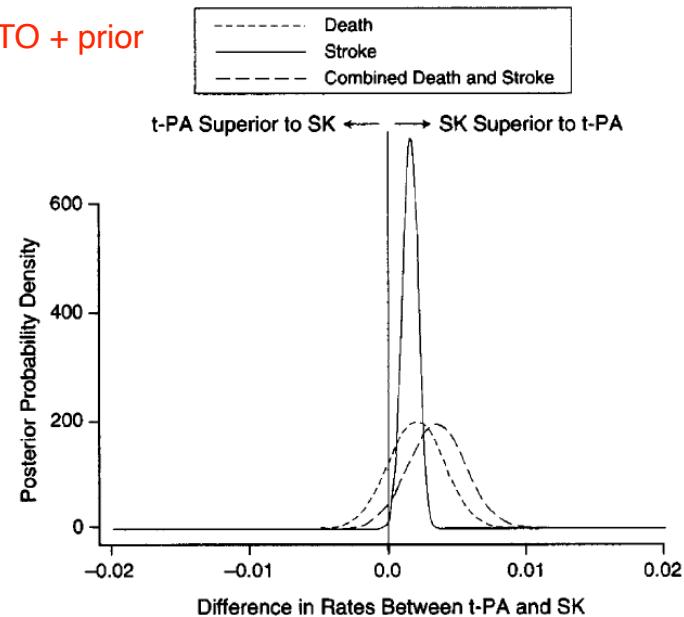
Prior Belief in GISSI-2 and ISIS-3, %	Probability of t-PA Mortality Higher Than SK Mortality	Probability of t-PA Net Clinical Benefit Greater Than SK Benefit	Probability of t-PA Net Clinical Benefit Greater Than SK Benefit by at Least 1%
100	.17	.05	<.001
50	.44	.24	<.001
10	.98	.94	.03
0	.999	.998	.36

*See footnote to Table 1 for expansions of abbreviations. Net clinical benefit is the combined death and stroke rate.

GUSTO alone



GUSTO + prior



PARAGON-HF

The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812

OCTOBER 24, 2019

VOL. 381 NO. 17

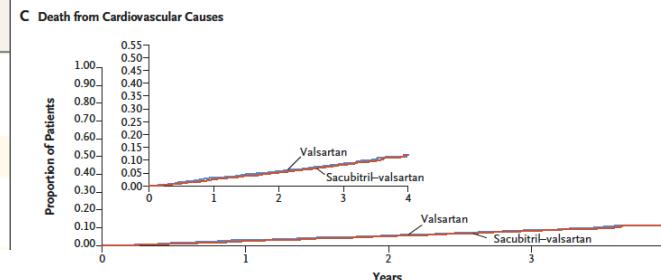
Angiotensin–Neprilysin Inhibition in Heart Failure with Preserved Ejection Fraction

METHODS

We randomly assigned 4822 patients with New York Heart Association (NYHA) class II to IV heart failure, ejection fraction of 45% or higher, elevated level of natriuretic peptides, and structural heart disease to receive sacubitril–valsartan (target dose, 97 mg of sacubitril with 103 mg of valsartan twice daily) or valsartan (target dose, 160 mg twice daily). The primary outcome was a composite of total hospitalizations for heart failure and death from cardiovascular causes. Primary outcome components, secondary outcomes (including NYHA class change, worsening renal function, and change in Kansas City Cardiomyopathy Questionnaire [KCCQ] clinical summary score [scale, 0 to 100, with higher scores indicating fewer symptoms and physical limitations]), and safety were also assessed.

Table 2. Primary and Secondary Outcomes.*

Outcome	Sacubitril–Valsartan (N=2407)	Valsartan (N=2389)	Ratio or Difference (95% CI)
Primary composite outcome and components			
Total hospitalizations for heart failure and death from cardiovascular causes†	P=0.06		RR, 0.87 (0.75–1.01)
Death from cardiovascular causes — no. (%)	204 (8.5)	212 (8.9)	HR, 0.95 (0.79–1.16)



CONCLUSIONS

Sacubitril–valsartan did not result in a significantly lower rate of total hospitalizations for heart failure and death from cardiovascular causes among patients with heart failure and an ejection fraction of 45% or higher. (Funded by Novartis; PARAGON-HF Clinical-

TEXT

The present data suggest that patients with a mildly reduced ejection fraction may have a response to sacubitril–valsartan...

PARAGON-HF - Frequentists being subjective

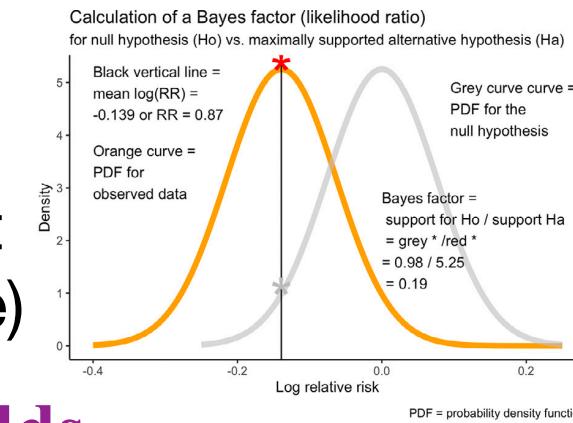


With this new proposed indication, sacubitril/valsartan would provide a safe and effective treatment for a broader range of HF patients, including those with HFpEF, a prevalent, progressive and debilitating condition with no approved treatment option.

- FDA - “biological plausibility” “ $p =0.06$ ” “post hoc analyses”
- IOW, it should work, it came close to working in the study and if we change the analyses it does work!
- But what, if any, was the quantitative reasoning behind this decision? How strong is the evidence? Can it be easily and accurately quantified? Do coherent actions follow?
- Answers to these questions are impossible with frequentist approach

PARAGON-HF - Bayesian inference

- 1° outcome RR 0.87, 95% CI 0.75-1.01, P = 0.06 ($z=1.89$)
- Minimum Bayes factor $= e^{-Z^2/2} = 1 / 5.85 = .17$
- H_0 (no effect) supported at 17% of the best supported H_a (p value over estimates evidence)



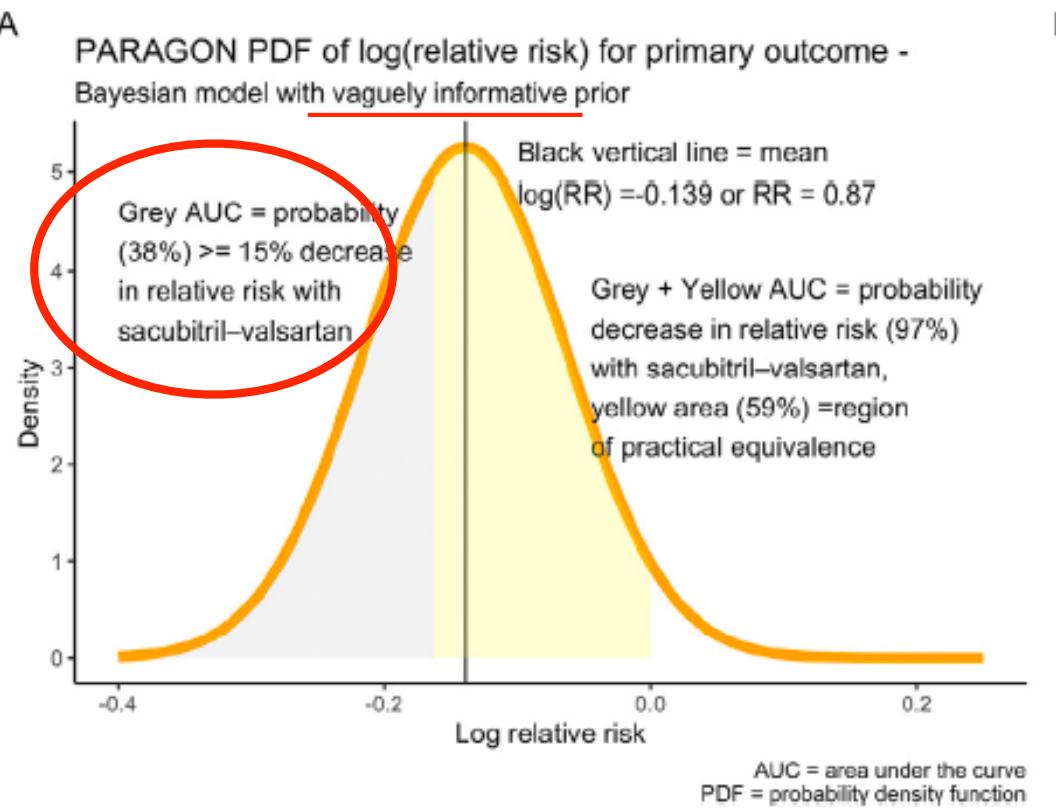
Posterior odds = likelihood ratio * Prior odds

- So if prior was coin toss, posterior odds = .17 or posterior probability of no effect = 14%
- Many failed HFrEF drugs, so if prior H_0 was 75%, then posterior probability of H_0 = 33% (no effect)
- For posterior probability of $H_0 < 5\%$ must believe that prior probability of no effect was $< 25\%$ (unrealistic given all previous failed drug trials for this condition)

$$\text{Probability} = \frac{\text{Posterior odds}}{1 + \text{Posterior odds}}$$

PARAGON-HF - Bayesian inference

- Bayesian analysis also helps differentiate statistical and clinical significance



- Low (**38%**) probability of clinical significance (as defined by PARAGON-HF investigators), robust to different priors
- This suggests best course of action is not approval but rather encouragement of further research to reduce the uncertainties and better define any potential benefits

STITCH (2011)

ORIGINAL ARTICLE

Coronary-Artery Bypass Surgery in Patients with Left Ventricular Dysfunction

METHODS

Between July 2002 and May 2007, a total of 1212 patients with an ejection fraction of 35% or less and coronary artery disease amenable to CABG were randomly assigned to medical therapy alone (602 patients) or medical therapy plus CABG (610 patients). The primary outcome was the rate of death from any cause. Major second-

Table 2. Study Outcomes.*

Outcome	Medical Therapy (N=602)	CABG (N=610)	Hazard Ratio with CABG (95% CI)	P Value†
	no. (%)			
Primary outcome: rate of death from any cause	244 (41)	218 (36)	0.86 (0.72–1.04)	0.12

FU 56 months

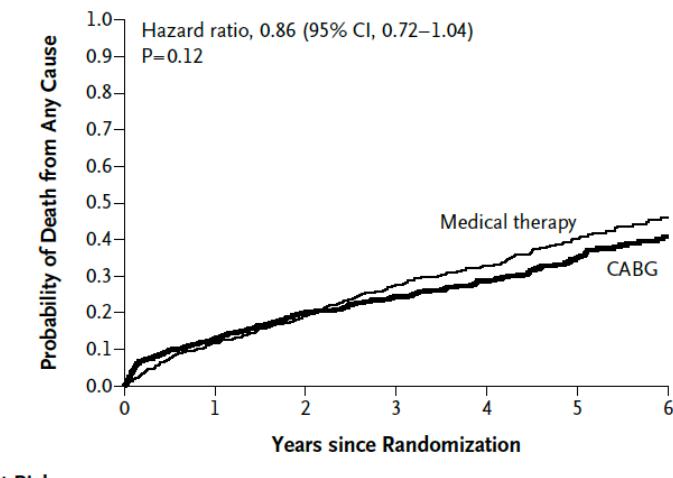


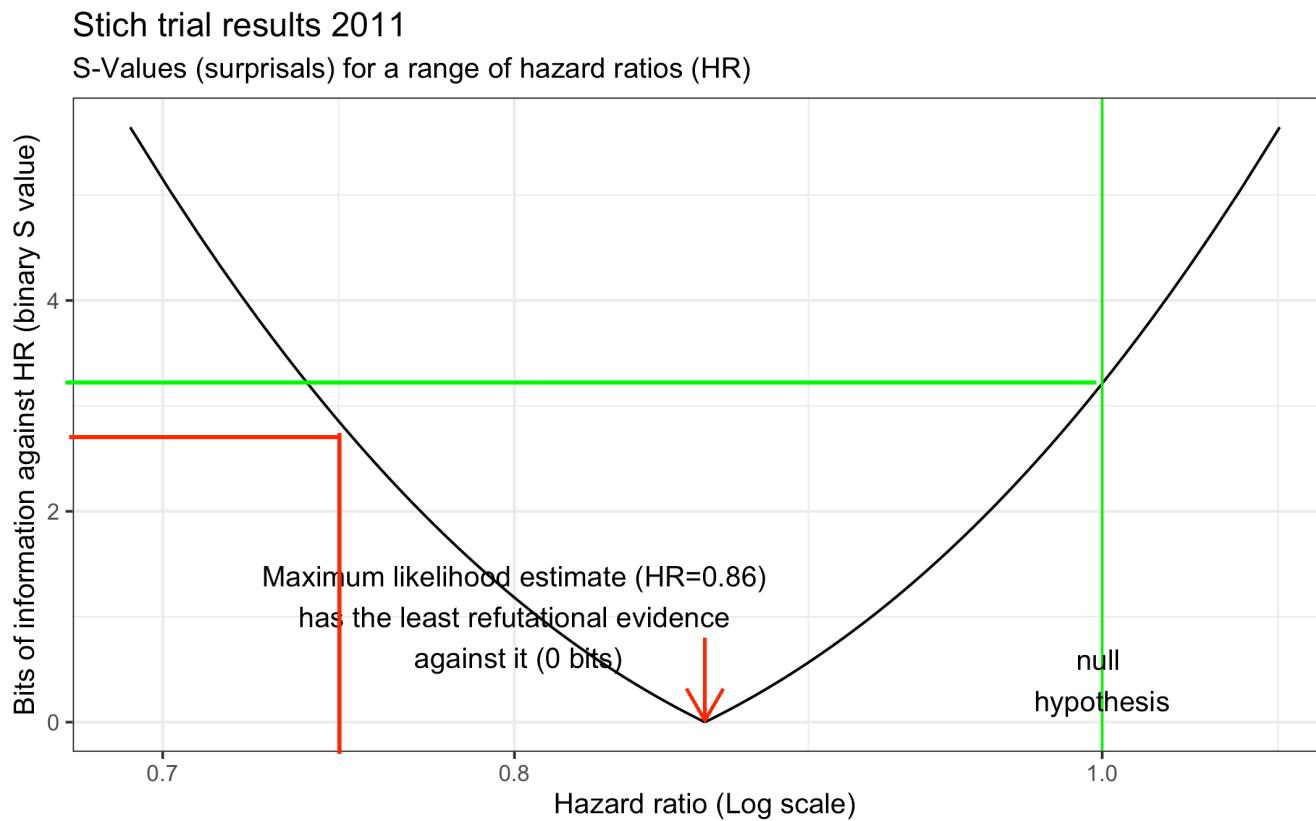
Figure 1. Kaplan-Meier Curves for the Probability of Death from Any Cause. CABG denotes coronary-artery bypass grafting.

CONCLUSIONS

In this randomized trial, there was no significant difference between medical therapy alone and medical therapy plus CABG with respect to the primary end point of death from any cause. Patients assigned to CABG, as compared with those assigned

The problem of nullism

- S value ($\log_2(1/P)$) measures strength of the evidence against not only H_0 but against any specific H_a - binary scale (# heads coin tosses)
- Consider the result $HR = 0.86$ (95% CI 0.72 - 1.04)



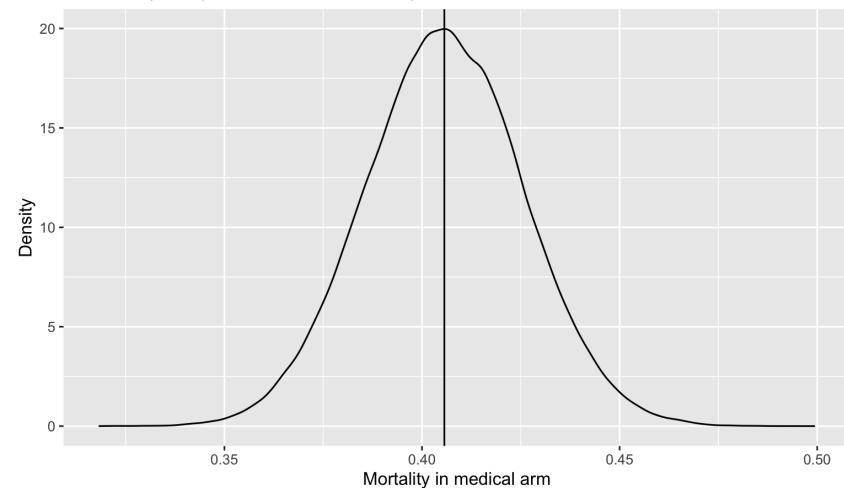
- < evidence against H_a = 25% decrease (red) than there is against H_0 hypothesis (green), which we have been told we should accept!

STITCH (2011) Bayesian perspective

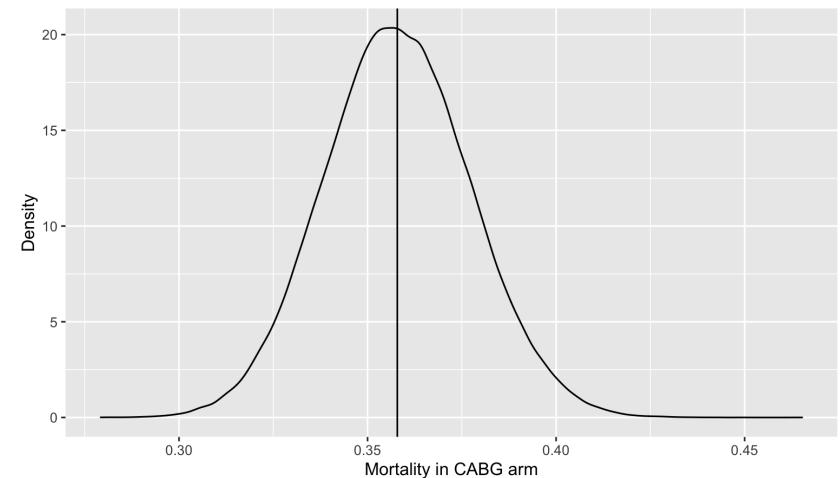
$$\text{posterior} = \frac{\text{prior} \cdot \text{likelihood}}{\text{normalizing constant}} \propto \text{prior} \cdot \text{likelihood}$$

- Likelihood \sim Binomial distribution of outcome in medical Tx arm & CABG arm (probability distribution for proportions)

STITCH (2011) PDF for total mortality in medical arm



STITCH (2011) PDF for total mortality in CABG arm

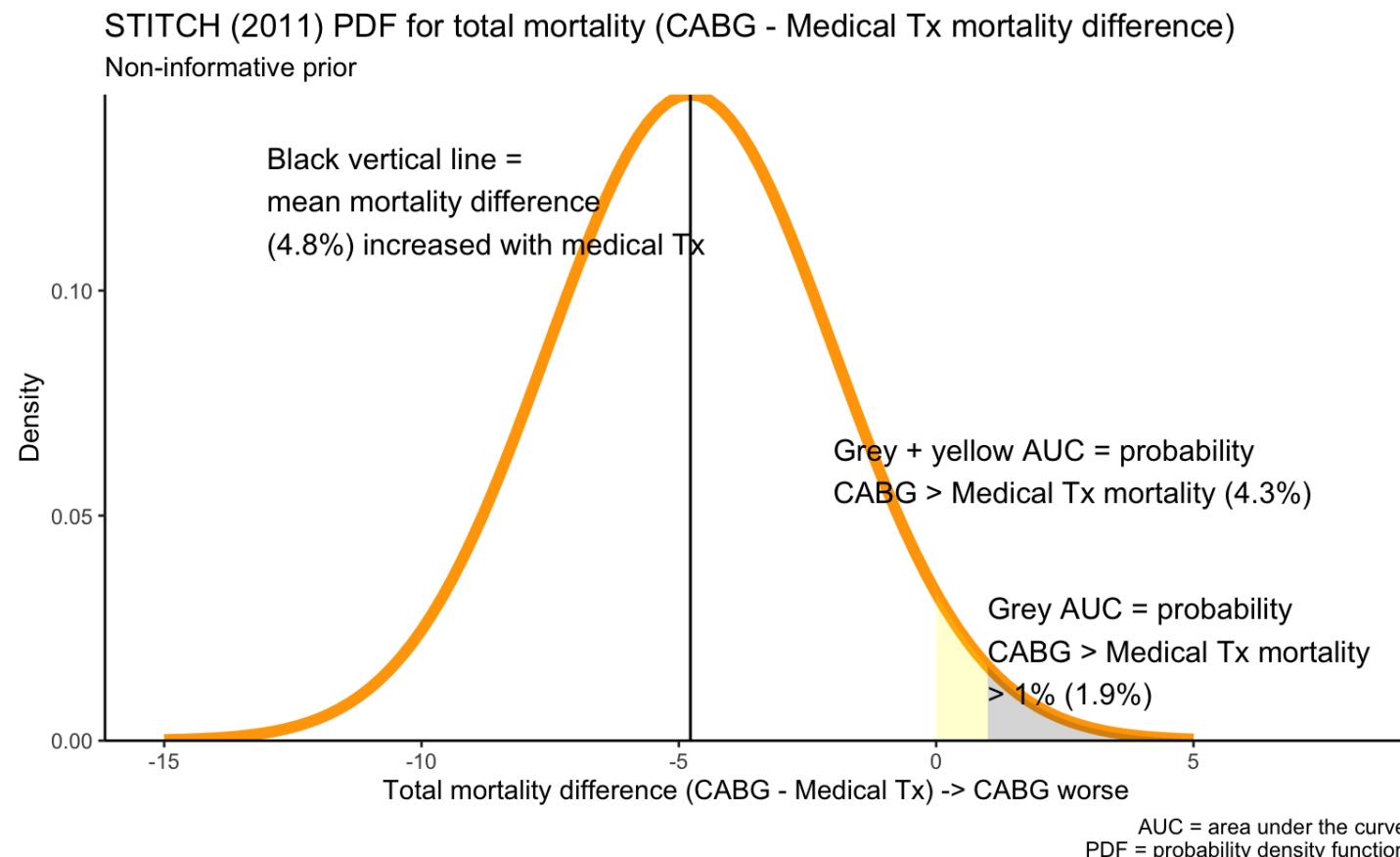


- Likelihood \sim Want probability for the difference ($\textcolor{red}{—}$) between the 2 distributions

STITCH (2011) Bayesian perspective

$$\text{posterior} = \frac{\text{prior} \cdot \text{likelihood}}{\text{normalizing constant}} \propto \text{prior} \cdot \text{likelihood}$$

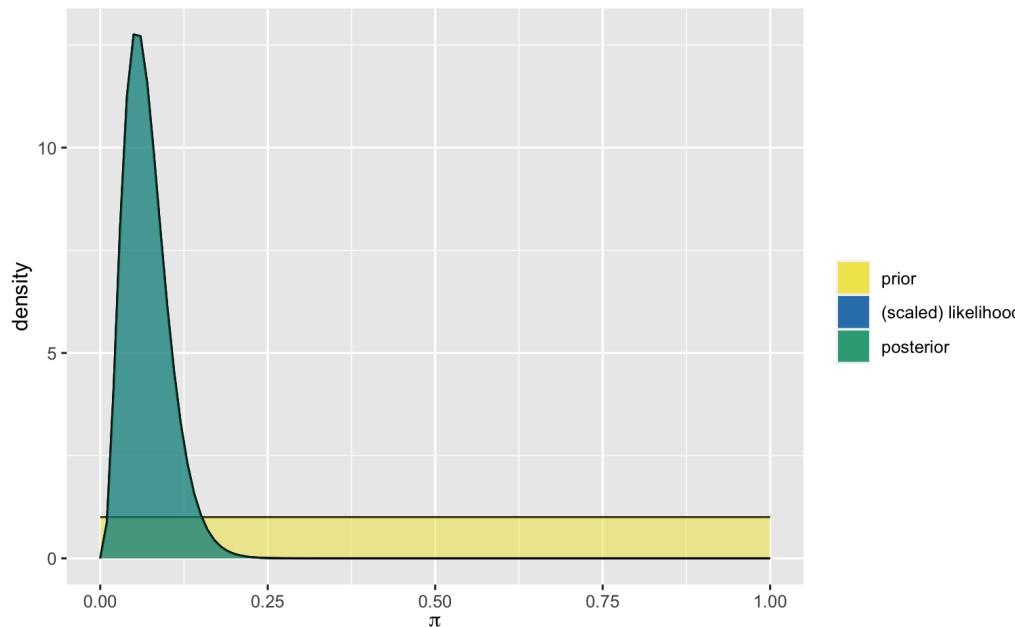
- Likelihood ~ Want probability for the difference ($\textcolor{red}{—}$) between the 2 distributions



STITCH (2011) Bayesian perspective

$$\text{posterior} = \frac{\text{prior} \cdot \text{likelihood}}{\text{normalizing constant}} \propto \text{prior} \cdot \text{likelihood}$$

- With a non-informative prior, posterior is dominated by data



- But prior data from RCT exists

THE NEW ENGLAND JOURNAL OF MEDICINE

June 27, 1985

A RANDOMIZED TRIAL OF CORONARY ARTERY BYPASS SURGERY

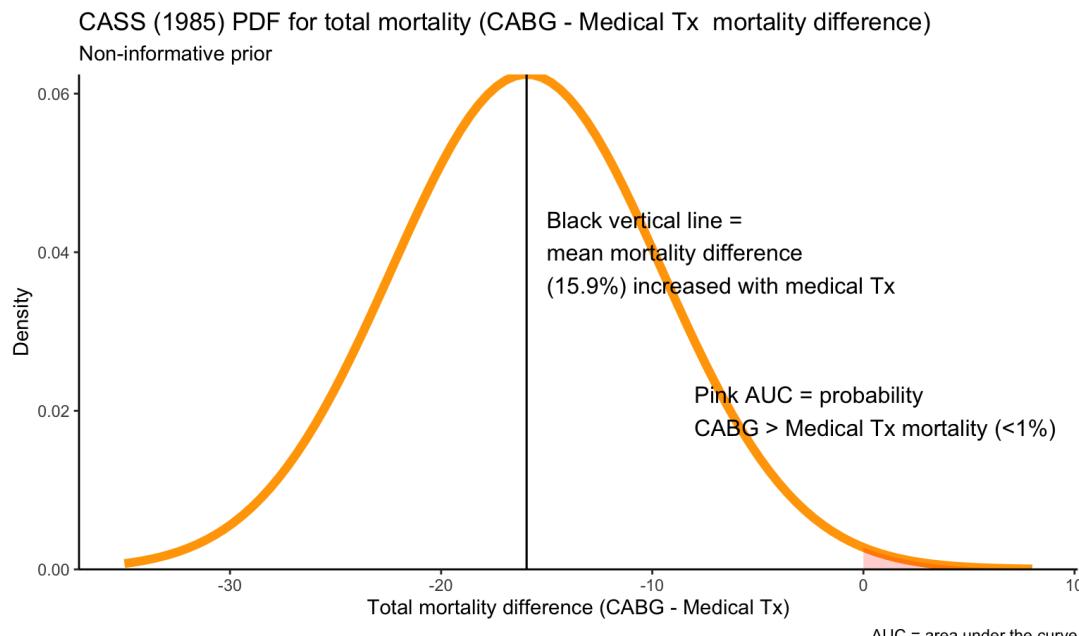
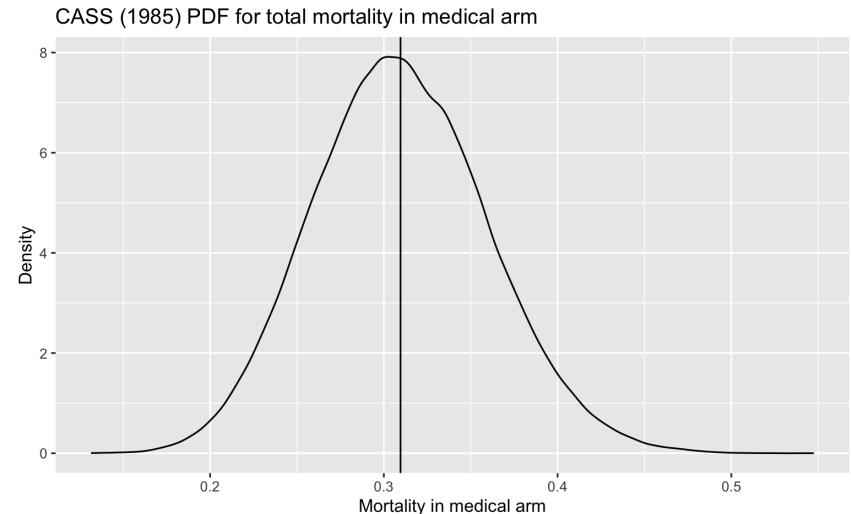
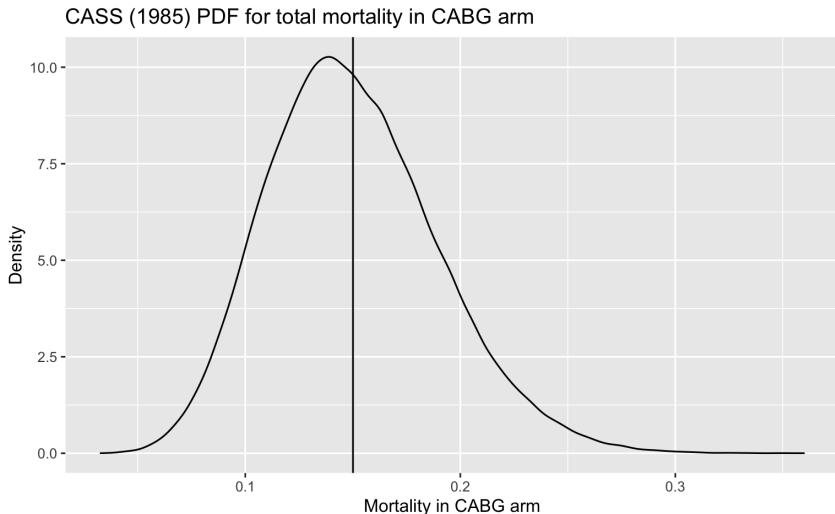
Survival of Patients with a Low Ejection Fraction

7 year mortality - 25 / 82 (medical 30%) versus 11 / 78 (CABG 14%)

STITCH (2011) Bayesian perspective

- Informative prior based on CASS (1985)

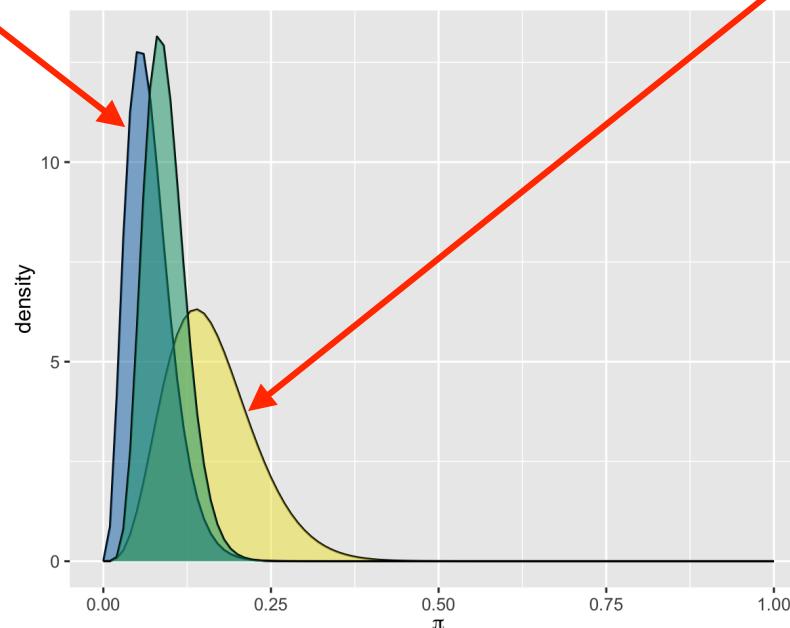
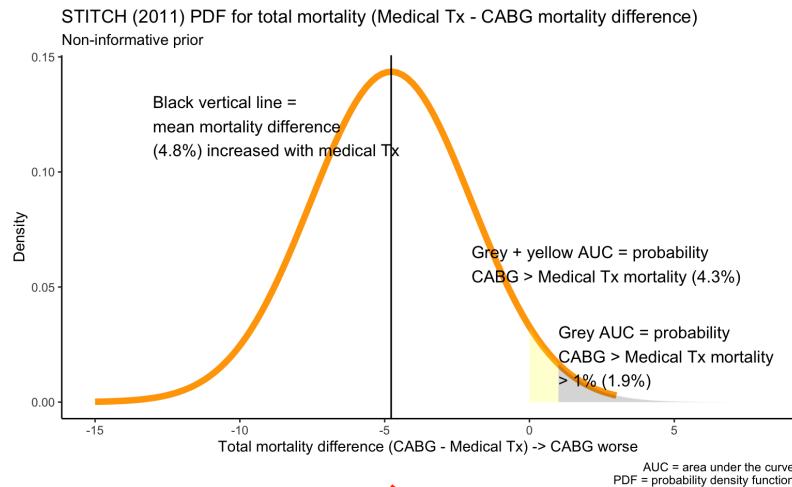
$$\text{posterior} = \frac{\text{prior} \cdot \text{likelihood}}{\text{normalizing constant}} \propto \text{prior} \cdot \text{likelihood}$$



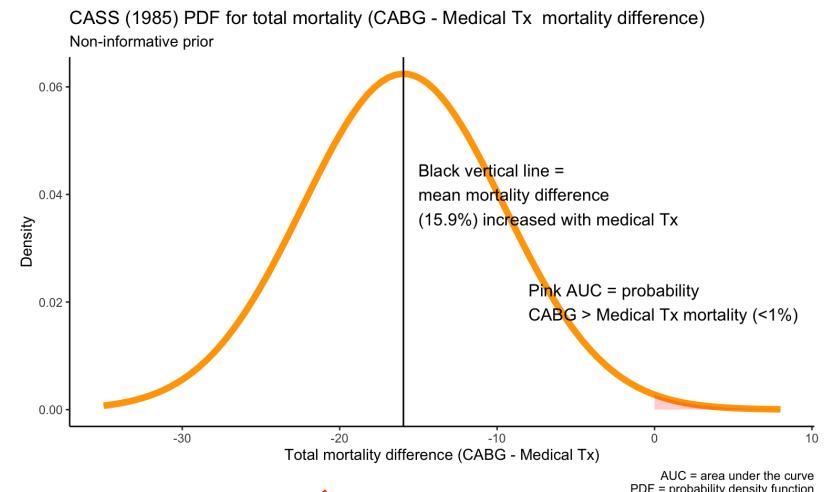
STITCH (2011) Bayesian perspective

STITCH data

$$\text{posterior} = \frac{\text{prior} \cdot \text{likelihood}}{\text{normalizing constant}} \propto \text{prior} \cdot \text{likelihood}$$

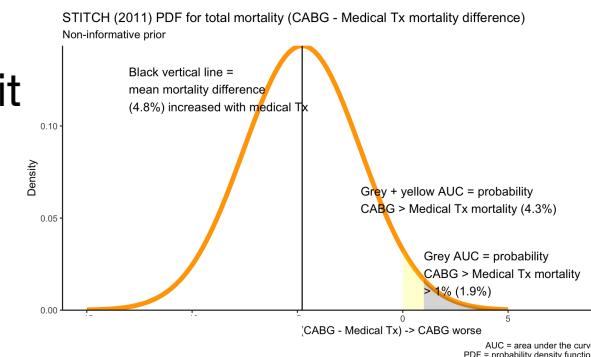
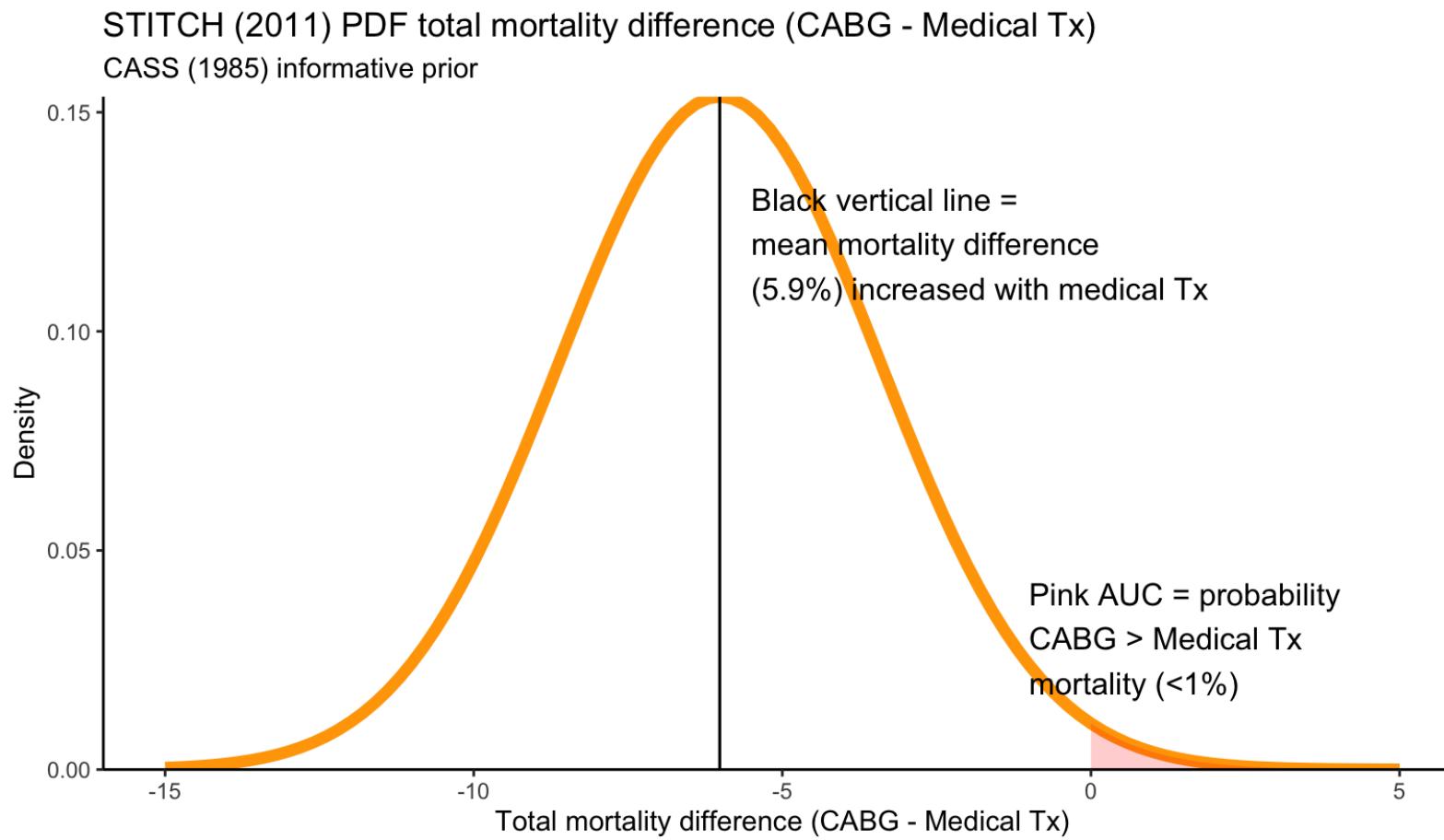


Informative prior - CASS (1985)



STITCH (2011)

- NEJM (2011) conclusion no benefit
- Posterior non-informative reasonably high probability CABG benefit
- Posterior with CASS(1985) informative prior, unequivocal large CABG benefit



STITCH (2016)

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Coronary-Artery Bypass Surgery in Patients with Ischemic Cardiomyopathy

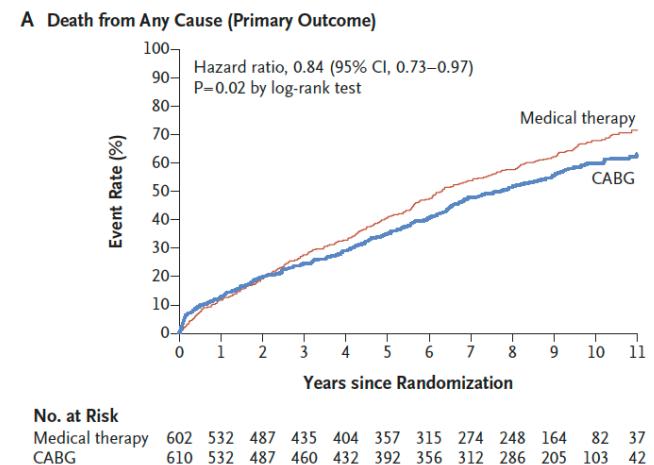
Table 2. Primary and Secondary Outcomes.

Outcomes	CABG Group (N=610)	Medical-Therapy Group (N=602)	Hazard Ratio (95% CI)*	P Value‡
	no. of patients (%)			
Primary outcome: death from any cause	359 (58.9)	398 (66.1)	0.84 (0.73–0.97)	0.02

FU 118 months

METHODS

From July 2002 to May 2007, a total of 1212 patients with an ejection fraction of 35% or less and coronary artery disease amenable to CABG were randomly assigned to undergo CABG plus medical therapy (CABG group, 610 patients) or medical therapy alone (medical-therapy group, 602 patients). The primary outcome was death from any cause. Major secondary outcomes included death from cardiovascular causes and death from any cause or hospitalization for cardiovascular causes. The median duration of follow-up, including the current extended-follow-up study, was 9.8 years.



CONCLUSIONS

In a cohort of patients with ischemic cardiomyopathy, the rates of death from any cause, death from cardiovascular causes, and death from any cause or hospitalization for cardiovascular causes were significantly lower over 10 years among patients who underwent CABG in addition to receiving medical therapy than among those who received medical therapy alone. (Funded by the National Institutes of

- Standard approach (2011 data) concluded no Δ between CABG and medical Tx in patients with LV dysfunction
- Bayesian approach (same 2011 data) with either non-informative or informative priors, high probability of a clinically meaningful benefit for CABG over medical Tx
- Bayesian result was confirmed with longer FU (2016 data)
- Bayesian approach -> CABG benefit offered 5 years earlier

RCT example

The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812

MAY 24, 2018

VOL. 378 NO. 21

Extracorporeal Membrane Oxygenation for Severe Acute Respiratory Distress Syndrome

RESULTS

At 60 days, 44 of 124 patients (35%) in the ECMO group and 57 of 125 (46%) in the control group had died (relative risk, 0.76; 95% confidence interval [CI], 0.55 to 1.04; **P=0.09**). Crossover to ECMO occurred a mean (\pm SD) of 6.5 ± 9.7 days after random-

CONCLUSIONS

Among patients with very severe ARDS, 60-day mortality was not significantly lower with ECMO than with a strategy of conventional mechanical ventilation that included ECMO as rescue therapy. (Funded by the Direction de la Recherche Clinique

TEXT

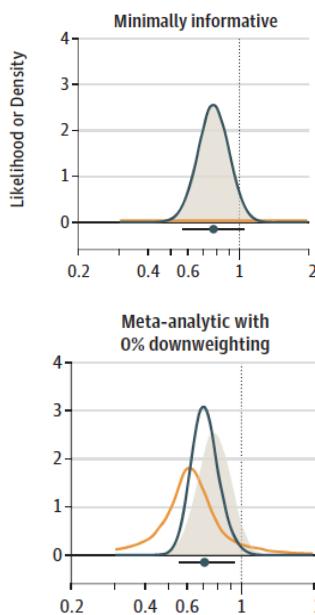
In conclusion, the analysis of the primary end point (mortality at 60 days) in our trial involving patients with very severe ARDS showed no significant benefit of early ECMO, as compared with a strategy of conventional mechanical ventilation, which included crossover to ECMO (used by 28% of the patients in the control group).

RCT - interpretation revisited

JAMA | Special Communication | CARING FOR THE CRITICALLY ILL PATIENT

Extracorporeal Membrane Oxygenation for Severe Acute Respiratory Distress Syndrome and Posterior Probability of Mortality Benefit in a Post Hoc Bayesian Analysis of a Randomized Clinical Trial

CONCLUSIONS AND RELEVANCE Post hoc Bayesian analysis of data from a randomized clinical trial of early extracorporeal membrane oxygenation compared with conventional lung-protective ventilation with the option for rescue extracorporeal membrane oxygenation among patients with very severe acute respiratory distress syndrome provides information about the posterior probability of mortality benefit under a broad set of assumptions that may help inform interpretation of the study findings.



Prior Belief	Posterior Median ARR, % (95% Credible Interval)	Posterior Probability That True ARR Is ≥ Specified Threshold, % ^a					
		2%	4%	6%	8%	10%	20%
Data-derived prior distribution							
No downweighting of previous studies	13.6 (2.9 to 20.5)	98	96	93	88	79	4
50% Downweighting of previous studies	12.8 (1.9 to 20.4)	97	95	91	83	72	3
75% Downweighting of previous studies	12.1 (1.1 to 20.3)	97	93	88	79	66	3

Other NEJM CV RCTs (with Bayesian benefits)

- Colchicine in patients with chronic coronary disease. N Engl J Med 2020;383:1838–47
- Efficacy and safety of low-dose colchicine after myocardial infarction. N Engl J Med 2019;381:2497–50
- Five-year outcomes after PCI or CABG for left main coronary disease. N Engl J Med 2019;381:1820–30.
- Rivaroxaban with or without aspirin in stable cardiovascular disease N Engl J Med 2017;377:1319-30
- Spironolactone for heart failure with preserved ejection fraction. N Engl J Med 2014;370:1383–92
- Ticagrelor versus clopidogrel in patients with acute coronary syndromes. N Engl J Med 2009;361:1045–57.
- An International Randomized Trial Comparing Four Thrombolytic Strategies for Acute Myocardial Infarction N Engl J Med 1993; 329:673-682

At the end of the day

- Frequentist Inference is Great For Doing Frequentist Inference
- Bayesian Inference is Great For Doing Bayesian Inference
- Mindless Frequentist Statistical Analysis is Harmful to Science
- Mindless Bayesian Statistical Analysis is Harmful to Science
- Frequentist & Bayesian Inference are often concordant
- But, sometimes they aren't & Bayesian Inference will add value

Of course, you won't know this unless
you do the Bayesian analysis

References

Frequentist statistics and inference

1. Ronald L. Wasserstein & Nicole A. Lazar (2016) The ASA Statement on p-Values: Context, Process, and Purpose, *The American Statistician*, 70:2, 129-133
2. Goodman, S.N. (1999a), "Toward Evidence-Based Medical Statistics 1: The P-Value Fallacy," *Annals of Internal Medicine*, 130, 995–1004
3. Greenland, S., and Poole, C. (2011), "Problems in Common Interpretations of Statistics in Scientific Articles, Expert Reports, and Testimony," *Jurimetrics*, 51, 113–129.
4. Brophy JM. Key issues in the statistical interpretation of randomized clinical trials. *Canadian Journal of Cardiology* (early e publication)

Bayesian statistics and inference

1. Goodman S. Toward Evidence-Based Medical Statistics. 2: The Bayes Factor. *Annals Int Med* 1999;130:1005-13.
2. Brophy JM. Bayesian analyses of cardiovascular trials—bringing added value to the table. *Canadian Journal of Cardiology* (early e publication)
3. Brophy JM. Bayesian Interpretation of the EXCEL Trial and Other Randomized Clinical Trials of Left Main Coronary Artery Revascularization. *JAMA Intern Med* 2020;180:986-92.
4. Brophy JM, Joseph L. Placing trials in context using Bayesian analysis. GUSTO revisited by Reverend Bayes. *JAMA* 1995;273:871-5.

General reading

1. Sharon Bertsch McGrayne. *The Theory That Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines, and Emerged Triumphant from Two Centuries of Controversy* - Google talk: <https://www.youtube.com/watch?v=8oD6eBkjF9o>

Quiz answers

#1 Show & main character - Big Bang Theory - Sheldon Cooper
What's cool - Bayes' Theorem on the whiteboard

#2 Choose A - hardcore frequentist
Choose B - latent bayesian

Thank you