A Bayesian Look at the Number Needed to Treat (NNT)

by

Lehana Thabane

McMaster University
Department Of Clinical Epidemiology and Biostatistics

Presented at the
Department of Mathematics and Statistics Seminar Series
McMaster University

October 2, 2002

Plan

- **Introduction**
  - ♣ Data Structure
  - ♣ Basic Measures of Risk/Benefit
  - ♣ Objectives
  - ♣ What is NNT?
    - ◇ Numerical Examples
    - ◇ Challenges with NNT
    - ◇ Applications of NNT in Health Research
  - ♣ Intro to Bayesian Approach
- **Posterior Distribution of NNT**
- **Some Benefits of Using the Bayesian Approach**
- **Pdf of NNT: Simulations**
    - ◇ Investigating the general behavior
    - ◇ Proposed "Bayesian Estimate"
    - ◇ Comparisons with other methods
- **Future Directions**
- **References**

1 Learning Objectives

- A Review of basic measures of clinical benefit
- A Review of the Bayesian Approach: what, how and when?
- What is NNT? Advantages and problems
- Derivation of the posterior distribution of NNT
- Insight on how to best estimate NNT?

2 Introduction

Table 1: Outcomes from a RCT

	Failure (Death)	Success /(Alive)
Intervention	a	b
Control	c	d

Table 2: Outcomes from a Case-Control Study

	Diseased (Cases)	Non-diseased (Controls)
Exposed	a	b
Not Exposed	c	d

2.1 Basic Measures

- Risk = Probability of death/disease

	Control	Intervention
Risk	p_1	p_2
Risk Estimate	$\frac{c}{c+d}$	$\frac{a}{a+b}$

♣ Risk Difference: RD= $p_1 - p_2$
♣ Relative Risk: RR= p_1/p_2
♣ Reduced Risk: $= \frac{p_1-p_2}{p_2} \times 100\% = (RR-1) \times 100\%$
♣ Odds: Odds=$\frac{p_i}{1-p_i}$
♣ Odds Ratio: OR $= \frac{p_1/(1-p_1)}{p_2/(1-p_2)}$

Primary Objectives

- To derive the posterior distribution of NNT

  ♣ Further our Understanding of the Distribution

  ♣ Provide better Estimation of NNT.

- **What is NNT?**

  ♣ First introduced by Laupacis, Sackett & Roberts (1988)

  ♣ $NNT = \frac{1}{|p_1 - p_2|}$

  ♣ *Interpretation:*

  ◇ If $p_1 - p_2 > 0$: The expected number of patients needed to treat to prevent one bad outcome or to get one benefit .

  ◇ If $p_1 - p_2 < 0$: The expected number of patients needed to treat to cause one bad outcome or to get one patient harmed.

  ♣ *Alternative Interpretation:* (JL Hutton. *JRSS Soc A* (2000); 163(3): 403-419)

  ◇ Average number of patients in the population 'needed to be treated' under new treatment to achieve one additional positive response (prevent one additional bad response) over the control.

Numerical Examples:

Example 1: Standard Treatment vs New Treatment

	Die	Survive	Total
Standard	11	55	67
New	1	62	63

- **Outcome Measure:** Proportion of deaths
- $p_1 = 11/67 = 0.164$ [Standard Treatment]
- $p_2 = 1/63 = 0.016$ [New Treatment]

  ♣ Relative Risk= $p_1/p_2 = 0.164/0.016 = 10.25$

  ♣ Risk Reduction= $p_1 - p_2 = 0.164 - 0.016 = 0.148$

  ♣ Odds Ratio= $\frac{p_1/(1-p_1)}{p_2/(1-p_2)} = \frac{0.164/0.836}{0.016/0.984} = 12.06$

  ♣ $NNT = \frac{1}{p_1-p_2} = 1/0.148 = 6.76$

Interpretations:

- $RR = 10.25$: The risk of death (probability of death) for people on standard treatment is about 10 times that for people on new treatment
- $RD = 14.8\%$: About 15% excess/additional risk (chances of death) for people on the standard compared to than those on new treatment
- $OR = 12.06$: About 12 times greater odds of death those on standard treatment than for those on new treatment
- $NNT = 6.76$: We need to treat about 7 patients to prevent one death

Example 2: Anti-epileptic Trial Data

	$\geq 50\%$ Reduction	$< 50\%$ Reduction	Total
Topiramate	8	15	23
Placebo	2	22	24

- **Source:** Sharief *et al* (1996): *Epilepsy Res* 25: 217-224
- Patients with at least one seizure/week during an 8-week baseline period.
- **Treatments:** Topiramate (400 mg/day) vs Placebo for 3 weeks; 8-week stabilization period.
- **Measure of effect:** At least 50% reduction in seizure rate at the end of treatment period.
- **Outcome Measure:** Proportion with at least 50% Reduction
- $p_1 = 8/23 = 0.8/23 = 0.35$ [Topiramate Treatment]
- $p_2 = 2/24 = 0.083$ [Placebo]

  ♣ Relative Risk$= p_1/p_2 = 0.35/0.083 = 4.17$

  ♣ Risk Reduction$= p_1 - p_2 = 0.35 - 0.083 = 0.26$

  ♣ Odds Ratio$= \frac{p_1/(1-p_1)}{p_2/(1-p_2)} = \frac{0.35/0.65}{0.083/0.917} = 5.87$

  ♣ $NNT = \frac{1}{p_1 - p_2} = 1/0.26 = 3.78$

  Interpretation: We need to treat about 4 patients to get one patient with at least 50 % reduction in seizure rate.

2.2 Applications of NNT in Health Research

- **Why use NNT?**

  ♣ Very Attractive measure to use from clinical perspective.

  ♣ Easier for Clinicians to interpret.

- **Applications**

  ♣ *Adverse outcomes:* Death, Stroke, adverse reaction, etc.

  ♣ *Beneficial Outcomes:* Improvement in quality of life or physical function, remission of symptoms, etc.

- **References:**

  ♣ *Screening:* Rembold CM. *BMJ*: 1998

  ♣ *Population & Disease Context:* Heller RF, Dobson AJ. *BMJ*: 2000

  ♣ *Clinical Medicine:* Sauve, Sauve, Sackett, *Clinical Research*: 1993

- **Challenges with NNT**

  ♣ Not defined at $p_1 - p_2 = 0$.

  ♣ Confidence Interval interpretation becomes difficult when (-L,U):

  ◇ Example: Suppose 95% CI for $p_1 - p_2$=(-0.25,0.25).

  ◇ Corresponding 95% CI for $NNT = (-\infty, -4]U[4, \infty)$.

  ♣ Overestimation: Jensen's Inequality: For a rv X and convex function $g(x)$, then

$$E(g(X)) \geq g(E(X))$$

Application to Estimation of NNT:

$$E(1/\hat{p}) \geq \frac{1}{E(\hat{p})} = \frac{1}{p}$$

  ♣ Invariance property of maximum likelihood estimation fails (Hutton JL (2000)).

Introduction to Bayesian Approach

- **Bayes Theorem:**

$$p(H|D) = \frac{P(D|H) \times P(H)}{P(D)}; \quad H = \text{Hypothesis}; \quad D = \text{Data}$$

- **Posterior**

$$p(\boldsymbol{\theta}|\mathbf{y}) = \frac{L(\boldsymbol{\theta}|\mathbf{y})p(\boldsymbol{\theta})}{p(\mathbf{y})} \quad [\textbf{Posterior Density}]$$

where

$$L(\boldsymbol{\theta}|\mathbf{y}) = \text{ likelihood function}$$
$$p(\theta) = \text{ prior density of } \boldsymbol{\theta}$$
$$p(\mathbf{y}) = \int_{\Omega} L(\boldsymbol{\theta}|\mathbf{y})p(\boldsymbol{\theta})d\boldsymbol{\theta}.$$

- **Inputs and Outputs of the Bayesian Analysis**

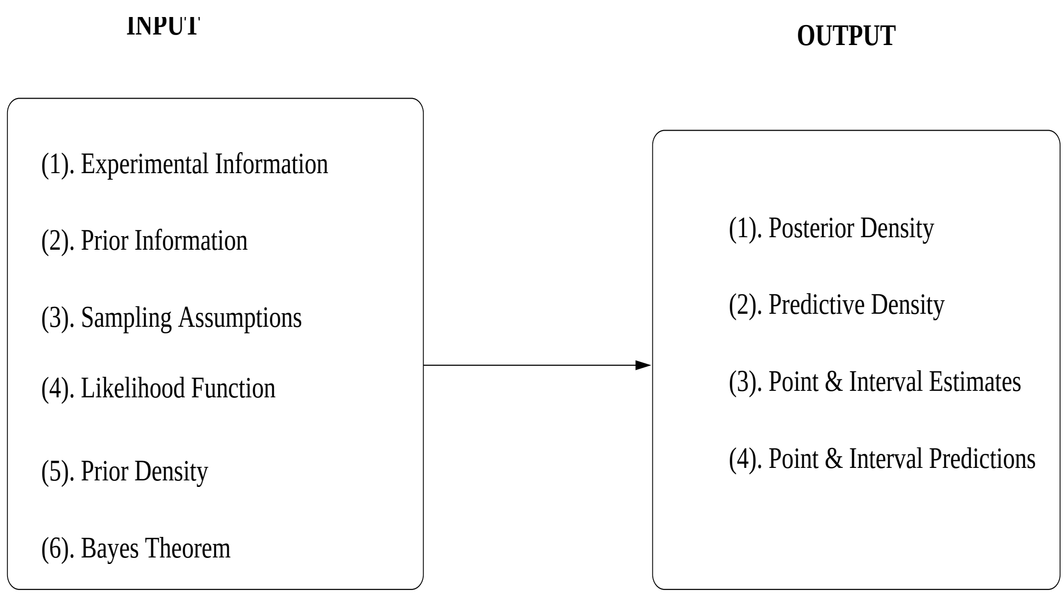

Figure 1: *Input and Output of Bayesian Analysis*

2.3 References on Bayesian Inference

- Introductory Level References
  - Berry DA (1996). *Statistics: A Bayesian Perspective*, Duxbury, London
  - Lee PM (1997). *Bayesian Statistics: An Introduction* 2nd Ed., Anorld, London
  - O'Hagan A (1988). *Probability: Methods and Measurement*, Chapman & Hall, London
  - Press SJ (1989). *Bayesian Statistics: Principles, Models and Applications*, Wiley, NY
- Intermediate→ Advanced
  - Bernardo JM and Smith AFM (1994). *Bayesian Theory*, Wiley, NY
  - O'Hagan A(1994). *Bayesian Inference*, Vol. 2B of "Kendall's Advanced Theory of Statistics", Arnold, London
- Bayesian Prediction (Introductory)
  - Geisser, S. (1993). *Predictive Inference: An Introduction*, Chapman and Hall, New York.
  - Aitchison, J. and Dunsmore, I.R. (1975). *Statistical Prediction Analysis*, Cambridge University Press, Cambridge.

3 Posterior Distribution of $p = p_1 - p_2$

- Likelihood function:

$$L(p_1, p_2|D) = \prod_{i=1}^{2} \binom{n_i}{x_i} p_i^{x_i} (1 - p_i)^{n_i - x_i} \tag{3.1}$$

- Prior Distribution: $p_i \sim \text{Beta}(\alpha_i, \beta_i)$
- Joint Posterior of (p_1, p_2)

$$f(p_1, p_2|D) = \prod_{i=1}^{2} \frac{\Gamma(x_i + \alpha_i)\,\Gamma(n_i - x_i + \beta_i)}{\Gamma(n_i + \alpha_i + \beta_i)} p_i^{x_i + \beta_i - 1} (1 - p_i)^{n_i - x_i + \beta_i - 1}. \tag{3.2}$$

- Mean and Variance of $p = p_1 - p_2$:

$$\begin{aligned} \mu_p &= E(p|D) = E(p_1|D) - E(p_2|D) \\ &= \frac{x_1 + \alpha_1}{n_1 + \alpha_1 + \beta_1} - \frac{x_2 + \alpha_2}{n_2 + \alpha_2 + \beta_2} \end{aligned}$$

$$\begin{aligned} \sigma_p^2 &= \text{Var}(p) = \sum_{i=1}^{2} \text{Var}(p_i|D) \\ &= \sum_{i=1}^{2} \frac{(x_i + \alpha_i)(n_i - x_i + \beta_i)}{(n_i + \alpha_i + \beta_i)^2 (n_i + \alpha_i + \beta_i + 1)}. \end{aligned}$$

References

1. Pham-Gia T. Value of the Beta Prior Information. *Commun Statist-Theory Meth* 23(8): 2175-95 (1994).

2. Pham-Gia T, Turkkan, N. Bayesian Analysis of the Difference of two proportions. *Commun Statist-Theory Meth* 22(6): 1755-71 (1993).

3. Geisser S. On Prior Distributions for Binary Trials (with discussions). *The American Statistian* 38: 244-51 (1984).

4 Asymptotic Posterior Distribution of p & NNT

- Asymptotic Posterior Distribution of p:

$$f(p|D) = \frac{1}{\sqrt{2\pi}\sigma_p} \exp\left\{-\frac{(p-\mu_p)^2}{2\sigma_p^2}\right\}. \tag{4.3}$$

- Asymptotic Posterior Distribution of NNT $= y = 1/p$

$$f(y|D) = \frac{1}{\sqrt{2\pi}\sigma_p y^2} \exp\left\{-\frac{(1/y-\mu_p)^2}{2\sigma_p^2}\right\}. \tag{4.4}$$

- Generalized Inverse Normal Family: Robert (1991), Johnson *at al* (1995, p.171)

$$p(y) = \frac{K(\alpha,\mu,\sigma)}{|y|^\alpha} \exp\left\{-\frac{(1/y-\mu)}{2\sigma^2}\right\}, \quad \alpha > 0, \sigma > 0 \;\; -\infty < \mu, y < \infty, \tag{4.5}$$

  ♣ k^{th} moment exists only if $\alpha > k+1$

  ♣ Modes at

$$y_1 = -\frac{\mu + \sqrt{\mu^2 + 4\alpha\sigma^2}}{2\alpha\sigma^2} \text{ and } y_2 = \frac{\sqrt{\mu^2+4\alpha\sigma^2} - \mu}{2\alpha\sigma^2}.$$

- From (4.4), the modes are

$$N\hat{N}T_1 = -\frac{\mu_p + \sqrt{\mu_p^2 + 8\sigma_p^2}}{4\sigma_p^2} \text{ and } N\hat{N}T_2 = \frac{\sqrt{\mu_p^2 + 8\sigma_p^2} - \mu_p}{4\sigma_p^2}. \tag{4.6}$$

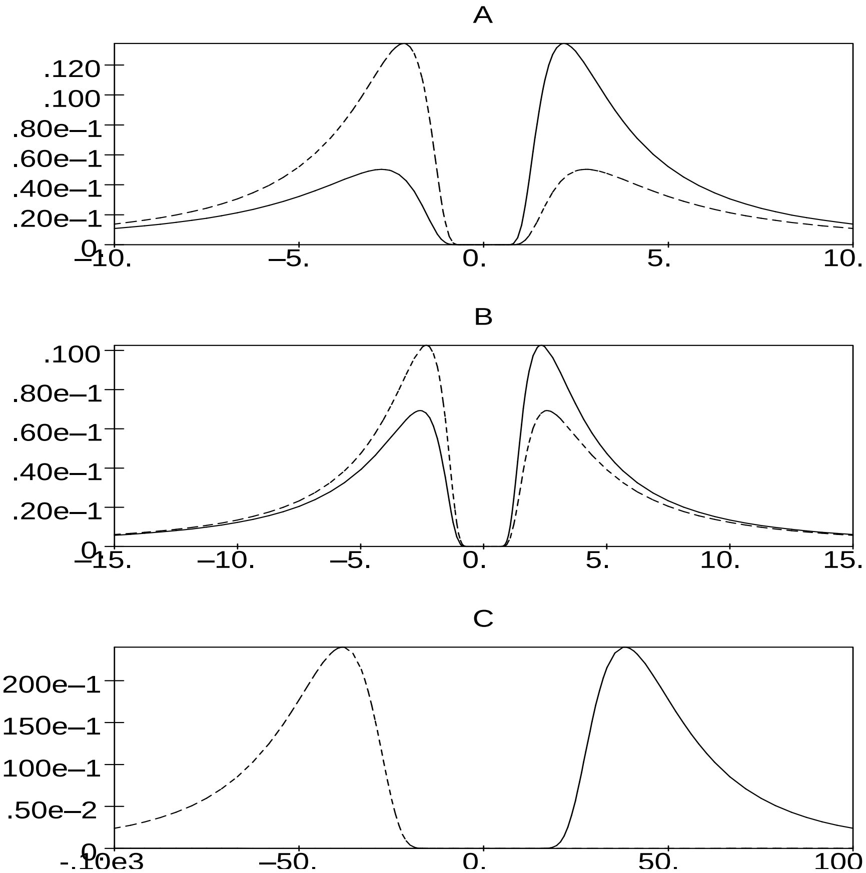

Figure 2: Posterior Distribution of NNT for A : $\mu_p = -0.1, 0.1, ;$ $\sigma_p = \sqrt{1/12} B$: $\mu_p = -0$ for:

Tf 365.76 m 432.34 -3674 95 Tz /F2 11.95 Tf 408.94 -30 TD[(1)1 T T Tf TD T NNT Tf

5 Benefits of Adopting Bayesian Approach

- Intuitive interpretation of credible intervals
- Uncertainly about NNT expressed explicitly through its posterior density
- Posterior Odds of Needed to Treat at least k subjects

$$\text{Odds}(NNT \geq k) = \frac{\Phi\left(\frac{\frac{1}{k}-\mu_p}{\sigma_p}\right)}{1-\Phi\left(\frac{\frac{1}{k}-\mu_p}{\sigma_p}\right)}.$$

- More on advantages of Bayesian approach in health research:
  - Wingler RL. Why Bayesian Analysis hasn't cought on in healthcare decision making. *Int J Tech Assess Health Care* 17 (1):56-66 (2001)
  - Hornberger J. Introduction to Bayesian Reasoning. *Int J Tech Assess Health Care* 17 (1):9-16 (2001)

6 Simulations

6.1 Objectives

- To study the behaviour of the posterior distribution of NNT
- To compare the performance of the posterior mode (Bayesian "Estimator") with conventional Estimators
  1. Classical Estimator: $N\hat{N}T_a = (x_1/n_1 - x_2/n_2)^{-1}$
  2. Adjusted Estimator: $N\hat{N}T_b = \left(\frac{x_1+1}{n_1+2} - \frac{x_2+1}{n_2+2}\right)^{-1}$
  3. Posterior Mode: $N\hat{N}T_c$

6.2 Results: Behaviour of Posterior Distribution

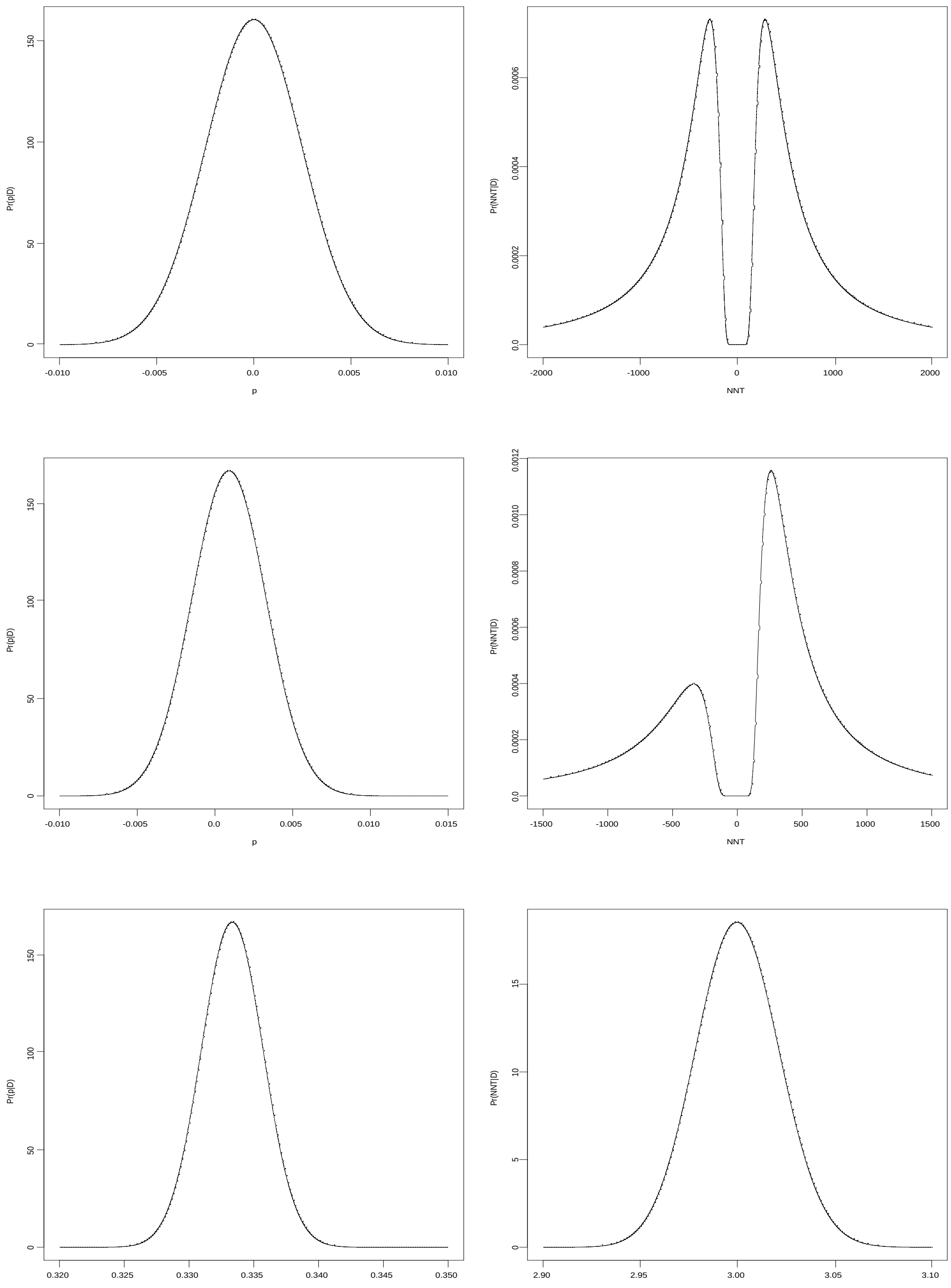

Figure 3: Posterior Distribution of p and NNT

6.3 Comparison of Estimators

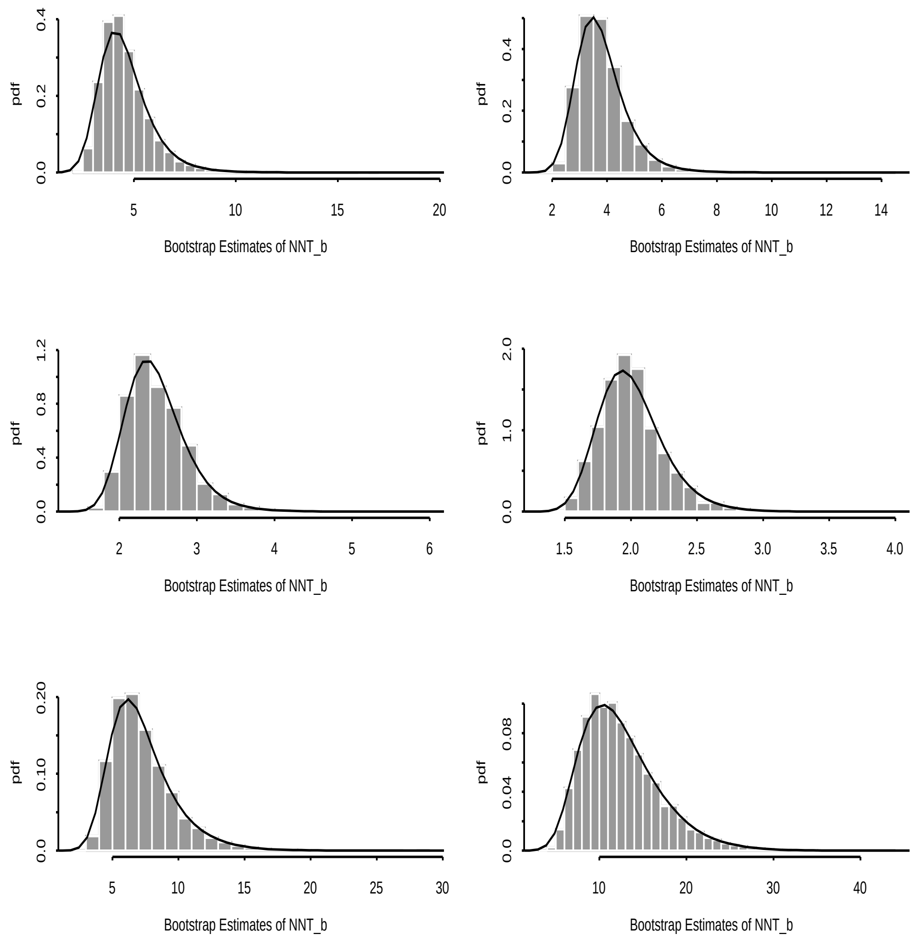

Figure 4: Distribution of Bootstrap Estimates based on NNT_b for $NNT = 5, 4, 2.5, 2, 6.67, 10$

Table 1: Average Percentage Error based on 100,000 simulations

			Average % Error= $\frac{\lvert Estimate - NNT \rvert}{NNT} \times 100$		
(p_1, p_2)	NNT	(n_1, n_2)	NNT_a	NNT_b	NNT_c
(0.8,0.1)	1.43	(100,100)	5.76	6.10	5.64
		(150,150)	4.69	4.87	4.63
		(250,250)	3.62	3.71	3.59
		(300,300)	3.29	3.35	3.26
(0.8,0.2)	1.67	(100,100)	7.62	7.94	7.22
		(150,150)	6.23	6.39	6.01
		(250,250)	4.80	4.88	4.70
		(300,300)	4.38	4.45	4.31
(0.8,0.3)	2	(100,100)	10.02	10.32	9.21
		(150,150)	8.11	8.28	7.67
		(250,250)	6.25	6.32	6.04
		(300,300)	5.65	5.71	5.50
(0.8,0.34)	2.5	(100,100)	13.40	13.78	11.70
		(150,150)	10.68	10.88	9.78
		(250,250)	8.18	8.27	7.77
		(300,300)	7.41	7.48	7.10
(0.8,0.45)	2.86	(100,100)	13.74	16.17	13.23
		(150,150)	12.47	12.66	11.12
		(250,250)	9.51	9.60	8.88
		(300,300)	8.57	8.65	8.13
(0.8,0.5)	3.33	(100,100)	19.12	19.63	15.10
		(150,150)	14.90	15.16	12.81
		(250,250)	11.17	11.29	10.23
		(300,300)	10.14	10.23	9.44

6.4 General Comments

1. The plots show that while the posterior distribution of p is nicely symmetric, that of NNT is not.
2. The posterior mode consistently gives the least average error percentages.
3. It out-performs the other conventional estimators if the support of the distribution lies entirely in the positive range, (*i.e.* if the probability of negative NNT is zero or very close to zero).

7 Future Directions

- The case of bimodality: The support of the distribution lies in both positive and negative axes.
- Meta-Analysis

References

[1] Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. *New Engl J Med* 318: 1728-33 (1988).

[2] Lesaffre E, Pledger G. A Note on the Number Needed to Treat. *Control Clinical Trials* 20: 439-47 (1999).

[3] Altman DG. Confidence Intervals for the number needed to treat. *BMJ* 317: 1309-12 (1998).

[4] Altman DG, Andersen PK. Calculating the number needed to treat where the outcome is time to an event. *BMJ* 319: 1492-95 (1999).

[5] Hutton, JL. Number needed to treat: properties and problems. *JRSS Ser A* 163: 403-15 (2000).

[6] Cates CJ. Simpson's paradox and the calculation of number needed to treat. *BMC Medical Research Methodology* 2:1 (2002) available at *http://www.biomedcentral.com/1471-2288/2/1*

[7] Altman DG, Deeks JJ. Meta-analysis, Simpson's paradox, and the number needed to treat. *BMC Medical Research Methedology* 2:3 (2002) available at *http://www.biomedcentral.com/1471-2288/2/3*

[8] Moore RA, Gavaghan DJ, Edwards JE, Wiffen W, McQuay HJ. Pooling data for Number Needed to Treat: no problems for apples. *BMC Medical Research Methedology* 2:2 (2002) available at *http://www.biomedcentral.com/1471-2288/2/2*

[9] Berger JO. *Statistical Decision Theory and Bayesian Analysis* 2nd Edition, Springer-Verlag, NY; 1985

[10] Bernardo JM, Smith AFM. *Bayesian theory*, Chichester, England: Wiley (1994)

[11] Berry DA. *Statistics: A Bayesian perspective.* Belmost, CA: Duxbury Press (1996)

[12] O'Hagan A. *Kendall's Advanced Theory of Statistics: Bayesian Inference*, Vol. 2B, Halsted Press, NY; 1996

[13] Pham-Gia T. Value of the Beta Prior Information. *Commun Statist-Theory Meth* 23(8): 2175-95 (1994).

[14] Pham-Gia T, Turkkan, N. Bayesian Analysis of the Difference of two proportions. *Commun Statist-Theory Meth* 22(6): 1755-71 (1993).

[15] Geisser S. On Prior Distributions for Binary Trials (with discussions). *The American Statistian* 38: 244-51 (1984).

[16] Bender R. Calculating Confidence Intervals for the Number Needed to Treat. *Control Clinical Trials* 22: 102-10 (2001).

[17] Walter SD. Number needed to treat (NNT): estimation of a measure of clinical benefit. *Stat Med* 20: 3947-62 (2001).

[18] Cook RJ, Sackett, DL. The number needed to treat: a clinically useful measure of effect. *MBJ* 310: 452-54 (1995).

[19] Sackett DL. On some clinically useful measures of the effects of treatment. *Evidenced-Based Med* 1: 37-38 (1996)

[20] Chatellier G, Zapletal E,*et al*. The number needed to treat: A clinically useful nomogram in its context. *BMJ* 12: 426-29 (1996).

[21] Newcombe RG. Confidence intervals for the number needed to treat: Absolute risk reduction is less likely to be misunderstood. *BMJ* 318:1765 (1999)

[22] North D. Number needed to treat: Absolute risk reduction may be easier for patients to understand. *BMJ* 310: 1269 (1995)

[23] Pickin M, Nicholl J. Number who benefit per unit of treatment may be a more appropriate measure. *BMJ* 310: 1270 (1995).

[24] Agresti A, Caffo B. Simple and effective confidence intervals for proportions and differences of proportions result from adding two successes and two failures. *The American Statistician* 54:280-88 (2000).

[25] Schouten HJA. Simple and effective confidence intervals for the number needed to treat. *Control Clinical Trials* 23: 100-02 (2002).

[26] Robert C. Generalized inverse normal distributions. *Statist Prob Letters* 11: 37-41 (1991).

[27] Johnson NL, Kotz S, Balakrishnan N. *Continuous Univariate Distributions, Volume 1*, 2nd ed., John Wiley & Sons, New York, 1995.

[28] Robert CP, Casella G. *Monte Carlo Statistical Methods*, Springer, NY; 1999

[29] Wingler RL. Why Bayesian Analysis hasn't cought on in healthcare decision making. *Int J Tech Assess Health Care* 17 (1):56-66 (2001)

[30] Hornberger J. Introduction to Bayesian Reasoning. *Int J Tech Assess Health Care* 17 (1):9-16 (2001)