A mixed approach for proving non-inferiority with respect to binary endpoints

Valentin Rousson and Burkhardt Seifert

Biostatistics Unit
Institute for Social and Preventive Medicine
University of Zurich

`rousson@ifspm.uzh.ch`

Berne, ROeS 2007

Comparing treatments in a randomized clinical trial (RCT)

In a RCT, a new treatment should be

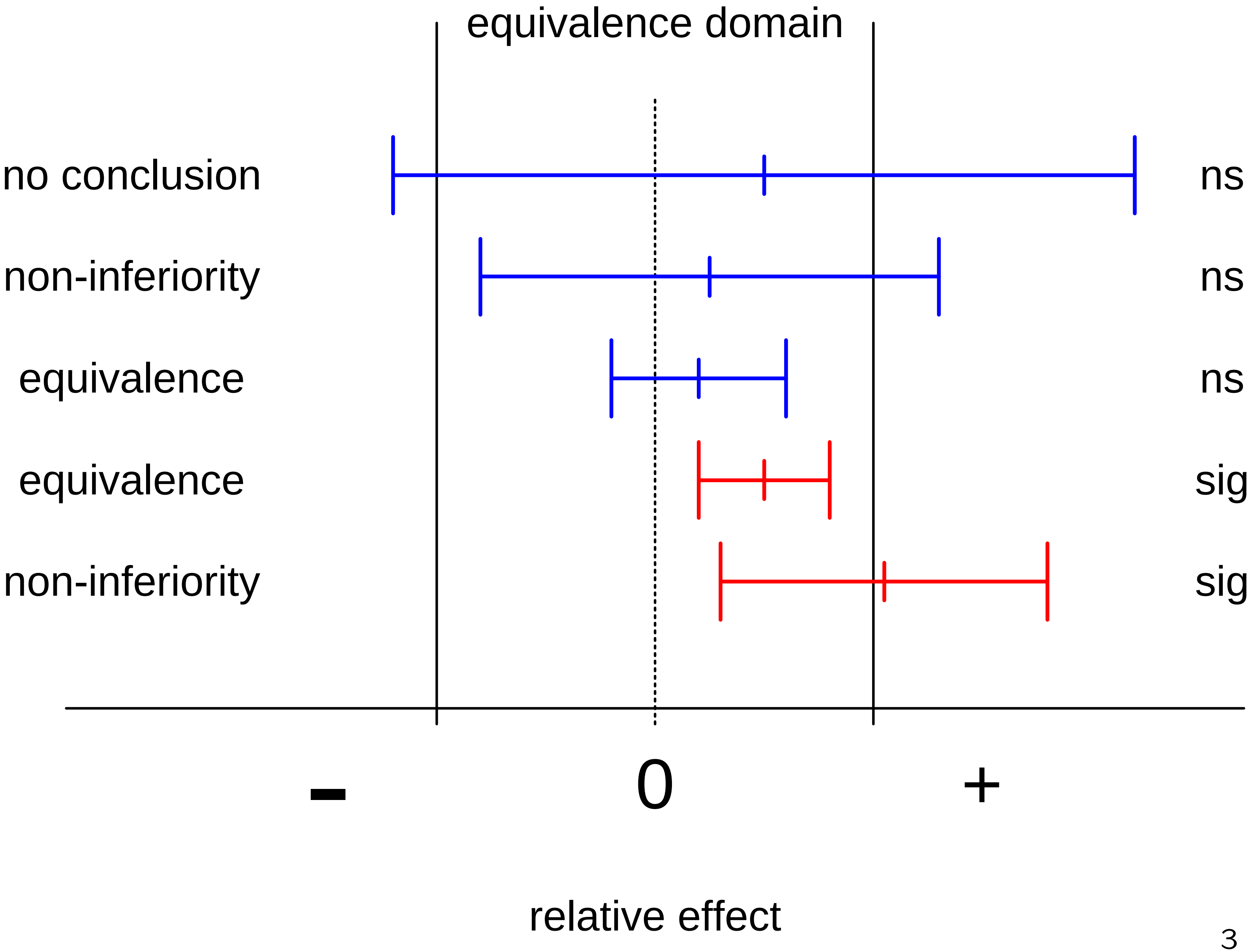
- Significantly better than a placebo
- Statistically non-inferior to an established treatment (active control)

Significantly better: Better (beyond chance)

Statistically non-inferior: Not much worse (beyond chance)

Can be assessed by comparing the left bound of a 95% CI for the relative effect of the new treatment with a non-inferiority margin (see e.g. Blackwelder, *Controlled Clinical Trials*, 1982)

Non-inferiority vs. equivalence testing



Measuring a relative effect for a binary endpoint

Let p_X be the proportion of success for the active control

Let p_Y be the proportion of success for the new treatment

These are measures of “absolute effects”

To measure a relative effect, one should compare both proportions

Candidates are:

- The difference $D = p_Y - p_X$
- The ratio $R = \frac{p_Y}{p_X}$
- The odds-ratio $OR = \frac{p_Y(1-p_X)}{p_X(1-p_Y)}$

Possible values

	D	R	OR
General	$(-1; 1)$	$(0; +\infty)$	$(0; +\infty)$
Given p_X	$(-p_X; 1 - p_X)$	$(0; 1/p_X)$	$(0; +\infty)$
$p_X = 0.2$	$(-0.2; 0.8)$	$(0; 5)$	$(0; +\infty)$
$p_X = 0.5$	$(-0.5; 0.5)$	$(0; 2)$	$(0; +\infty)$
$p_X = 0.8$	$(-0.8; 0.2)$	$(0; 1.25)$	$(0; +\infty)$

⇒ Domain of possible values of OR independent from the absolute efficacy of treatments (see e.g. Garrett, *Statistics in Medicine*, 2003)

⇒ The non-inferiority margin can be chosen independently from the absolute efficacy of treatments (for example $\varepsilon_{OR} = 0.5$)

Proportion of success and proportion of failure

Comparing proportions of failure $q_X = 1 - p_X$ and $q_Y = 1 - p_Y$ yields

- $D' = q_X - q_Y = (1 - p_X) - (1 - p_Y) = D$
- $R' = \frac{q_X}{q_Y} = \frac{1-p_X}{1-p_Y} \neq R$
- $OR' = \frac{q_X(1-q_Y)}{q_Y(1-q_X)} = \frac{(1-p_X)p_Y}{(1-p_Y)p_X} = OR$

Example: $p_X = 0.8$, $p_Y = 0.9$, $q_X = 0.2$ and $q_Y = 0.1$

- $D = D' = 0.1$
- $R = 1.125$ and $R' = 2$
- $OR = OR' = 2.25$

Relationship between D and OR

- OR and D are not one-to-one related

Examples:

If $p_X = 0.8$, $OR = 0.5$ corresponds to $D = -0.13$ ($p_Y = 0.67$)

If $p_X = 0.5$, $OR = 0.5$ corresponds to $D = -0.17$ ($p_Y = 0.33$)

- But given p_X, they are related as

$$D = \frac{p_X(1 - p_X)(OR - 1)}{1 + p_X(OR - 1)}$$

Example: If one knew that $p_X = 0.8$, it is the same to show $OR \geq 0.5$ than to show $D \geq -0.13$

Result: It is statistically easier to show the latter than to show the former if $p_X = p_Y$ is large!

Planning a non-inferiority study

- One usually assumes $OR = 1$ and aims to show $OR \geq \varepsilon_{OR}$

Assuming a value for $p_X = p_Y$, one should enroll per group

$$n_{OR} = \frac{2(z_\alpha + z_\beta)^2}{p_X(1 - p_X)(\log(\varepsilon_{OR}))^2}$$

- But given p_X, one may equivalently assume $D = 0$ and aims to show $D \geq \varepsilon_D$ with

$$\varepsilon_D = \frac{p_X(1 - p_X)(\varepsilon_{OR} - 1)}{1 + p_X(\varepsilon_{OR} - 1)}$$

Here, one should enroll per group

$$n_D = \frac{2(z_\alpha + z_\beta)^2 p_X(1 - p_X)}{\varepsilon_D^2}$$

Result: $n_D \leq n_{OR}$ if $p_X \geq \frac{1}{1-\varepsilon_{OR}} + \frac{1}{\log(\varepsilon_{OR})}$ (Rousson and Seifert, 2007)

Sample size comparison (n_{OR}/n_D)

		ε_{OR}			
$1-\beta$	$p_X = p_Y$	0.43	0.5	0.55	0.8
80%	0.5	89/99	131/142 (139)	176/187	1262/1272
	0.6	92/88	137/129 (126)	184/173	1314/1267
	0.7	105/84	156/127 (126)	210/174	1502/1383
	0.8	138/90	205/142 (145)	275/199	1971/1731
	0.9	245/128	364/212 (226)	489/305	3503/2932
	0.95	464/214	688/365 (401)	925/535	6638/5421
90%	0.5	119/133	175/190 (186)	236/250	1689/1703
	0.6	123/117	183/172 (169)	245/231	1759/1696
	0.7	141/112	209/170 (169)	280/232	2010/1851
	0.8	185/120	274/190 (193)	368/266	2638/2317
	0.9	328/171	486/283 (299)	654/409	4690/3926
	0.95	622/287	921/488 (531)	1238/717	8886/7257

(In parentheses are the n_D obtained using the formula of Farrington and Manning, *Statistics in Medicine*, 1990)

Mixed approach to show non-inferiority

1. Define the non-inferiority margin for OR, for example $\varepsilon_{OR} = 0.5$

2. Assume a proportion of success for the active control, for example $p_X = 0.8$

3. Using the one-to-one relationship above, calculate the corresponding non-inferiority margin for D, for example

$$\varepsilon_D = \frac{0.8(1-0.8)(0.5-1)}{1+0.8(0.5-1)} = -0.13$$

4. Calculate a 95% CI for D and check that its lower bound is larger than ε_D (-0.13 in that example)

Relationship between ε_D and ε_{OR} given p_X

	ε_{OR}			
p_X	0.43	0.5	0.55	0.8
0.5	−0.199	−0.167	−0.145	−0.056
0.55	−0.205	−0.171	−0.148	−0.056
0.6	−0.208	−0.171	−0.148	−0.055
0.65	−0.206	−0.169	−0.145	−0.052
0.7	−0.199	−0.162	−0.138	−0.049
0.75	−0.187	−0.150	−0.127	−0.044
0.8	−0.168	−0.133	−0.112	−0.038
0.85	−0.141	−0.111	−0.093	−0.031
0.9	−0.105	−0.082	−0.068	−0.022
0.95	−0.059	−0.045	−0.037	−0.012

Most of these margins lie between −0.1 and −0.2 and become wider for decreasing efficiencies, as recommended by the FDA (1997)

Approximate 95% CI for D and OR

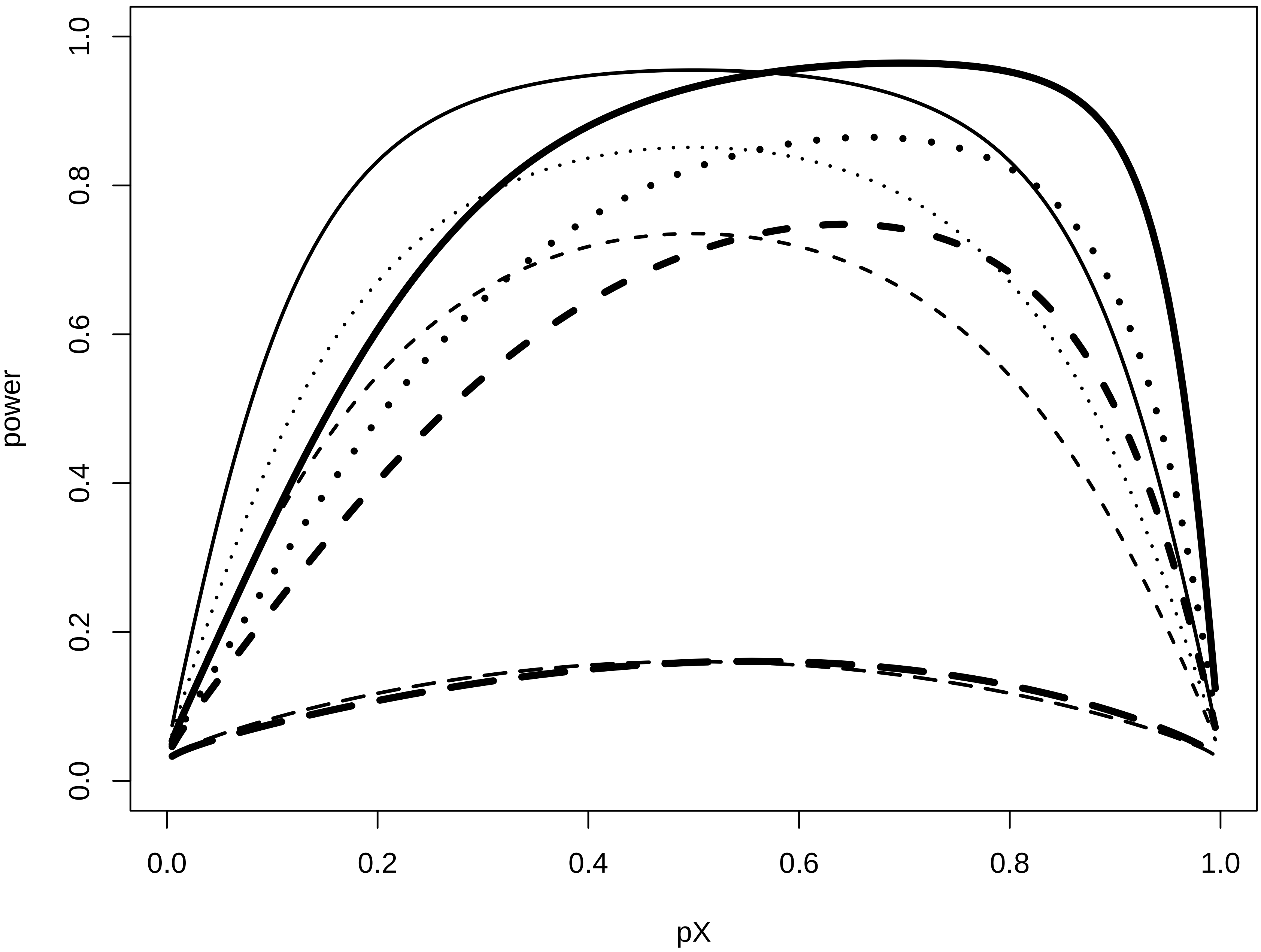
Let $\widehat{p}_X$ and $\widehat{p}_Y$ be empirical proportions of success from samples of size n, and let $\widehat{D} = \widehat{p}_Y - \widehat{p}_X$ and $\widehat{OR} = \widehat{p}_Y(1-\widehat{p}_X)/(\widehat{p}_X(1-\widehat{p}_Y))$

$$\widehat{D} \pm 1.96\sqrt{\frac{\widehat{p}_X(1-\widehat{p}_X)}{n} + \frac{\widehat{p}_Y(1-\widehat{p}_Y)}{n}} \tag{1}$$

$$\log(\widehat{OR}) \pm 1.96\sqrt{\frac{1}{\widehat{p}_X(1-\widehat{p}_X)n} + \frac{1}{\widehat{p}_Y(1-\widehat{p}_Y)n}} \tag{2}$$

- Conventional OR approach: Uses ε_{OR} and (2)
- Mixed approach: Uses $\varepsilon_D(\varepsilon_{OR}, p_X)$ and (1)

Power comparison ($n = 150$ per group)



Example

- Non-inferiority margin $\varepsilon_{OR} = 0.5$

  Assumed value for $p_X = 0.8$

  Corresponds to non-inferiority margin $\varepsilon_D = -0.13$

- Data: $n = 150$ patients per group

  $\widehat{p}_X = 125/150 = 0.833$ and $\widehat{p}_Y = 121/150 = 0.807$

  Corresponds to $\widehat{D} = -0.027$ and $\widehat{OR} = 0.83$

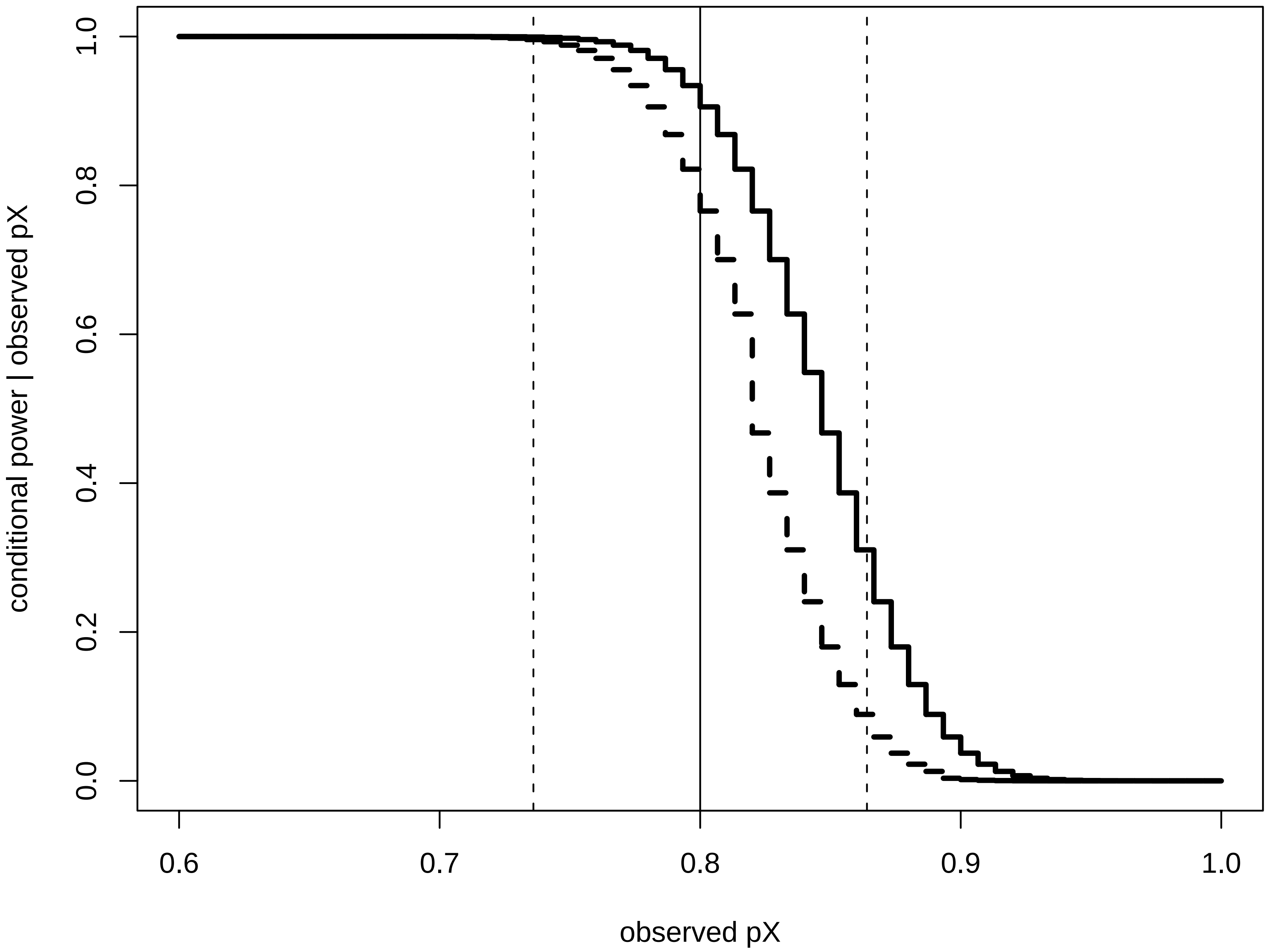
- Conventional OR approach:

  95% CI for OR is $[0.46; 1.50]$ ⇒ One fails to show non-inferiority

- Mixed approach:

  95% CI for D is $[-0.11; 0.06]$ ⇒ One has shown non-inferiority

Conditional power comparison ($n = 150$ per group)



Generalization to an ordinal enpoint

Let X and Y be outcomes of active control and new treatment on an ordinal scale (e.g. failure, moderate success, big success)

One usually measures relative effect by

$$\theta_1 = \Pr\{X < Y\} + \frac{\Pr\{X = Y\}}{2}$$

or

$$\theta_2 = \frac{\Pr\{X < Y\}}{1 - \Pr\{X = Y\}}$$

Equivalently, one can use

$$\theta_1^* = 2\theta_1 - 1 = \Pr\{X < Y\} - \Pr\{X > Y\}$$

or

$$\theta_2^* = \frac{\theta_2}{1 - \theta_2} = \frac{\Pr\{X < Y\}}{\Pr\{X > Y\}}$$

Generalization of D and OR

For a binary endpoint with p_X and p_Y, one has

$\Pr\{X < Y\} = (1 - p_X)p_Y$ and $\Pr\{X = Y\} = (1 - p_X)(1 - p_Y) + p_X p_Y$

Thus,

$$\theta_1^* = \Pr\{X < Y\} - \Pr\{X > Y\} = (1 - p_X)p_Y - (1 - p_Y)p_X = p_Y - p_X = D$$

and

$$\theta_2^* = \frac{\Pr\{X < Y\}}{\Pr\{X > Y\}} = \frac{(1 - p_X)p_Y}{(1 - p_Y)p_X} = OR$$

Remark: D and OR can further be generalized as

$$D = \Pi_c - \Pi_d$$ (Kendall's τ_a, *Biometrika*, 1938)

and

$$OR = \frac{\Pi_c}{\Pi_d}$$ (General OR of Agresti, *Biometrics*, 1980)

where Π_c and Π_d are probabilities of concordance/discordance

Generalization of the mixed approach

1. Define a non-inferiority margin ε_{OR}
2. Assume a distribution p_X for X (for example $p_X = (0.1, 0.2, 0.7)$)
3. Calculate a corresponding non-inferiority margin ε_D based on ε_{OR} and p_X (no unique solution, see Rousson and Seifert, 2007)
4. Calculate a 95% CI for D (see Edwardes, *Biometrics*, 1995) and check that its lower bound is larger than ε_D

Power comparison (3 possible outcomes, $n = 150$ per group)

	p_3							
p_2	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
0.1	72/53	84/68	90/80	92/87	92/92	**90/93**	**84/93**	**72/89**
0.2	84/69	91/81	94/89	95/93	94/95	**91/95**	**84/93**	
0.3	89/79	94/88	96/93	95/95	94/96	**89/94**		
0.4	91/86	94/92	95/95	95/95	92/94			
0.5	92/90	94/93	94/94	92/92				
0.6	90/91	91/92	89/88					
0.7	**84/88**	84/84						
0.8	**71/75**							