Optimal rejection regions for testing multiple binary endpoints in small samples

Robin Ristl and Martin Posch

Section for Medical Statistics, Center of Medical Statistics, Informatics and Intelligent Systems, Medical University of Vienna

Köln, June 2015

Possible aims of a study with multiple endpoints

(1) Show effect in all endpoints
- Reject all H_i
- Endpoints are co-primary

(2) Reject the global null-hypothesis of no effect in any EP
- Need not involve conclusions on individual EPs

(3) Identify efficacious endpoints
- Reject some H_i
- Multiple testing problem

(4) Show non-inferiority for some EP and superiority for at least one EP
- Might be considered as minimal condition for an improved treatment

What is the problem with small sample sizes?

- Asymptotic distibution may not reflect true distributions with sufficient accuracy
- Low precision of nuisance parameter estimates
- Limited model complexity/risk of overfitting
- Low power

Testing binary endpoints

- $H_0 : p_1 = p_2$ or equivalently $H_0 : OR = 1$
  - p_1, p_2 proportions in treatment and placebo group
  - n sample size per group
- Large sample sizes:
  - Asymptotic z-test
  - Under H_0, $Z = \sqrt{n}\frac{\hat{p}_T - \hat{p}_C}{\sqrt{2\hat{p}_0(1-\hat{p}_0)}} \xrightarrow{d} N(0,1)$
  - Uses central limit theorem and consistent estimate for p_0
- Small sample sizes:
  - Under H_0, $n\hat{p}_1, n\hat{p}_2 \sim Binom(n, p_0)$
  - Can find the exact distribution of $\hat{p}_1 - \hat{p}_2$ for given p_0
  - Difficulty: p_0 is unkown
- Conditional test
  - Table row margins are a sufficient statistic for p_0
  - Conditional on the margins the data do not depend on p_0

Fisher's Exact Test

- Null hypothesis H_0 : Odds ratio ≤ 1
- Test statistic T is the number of successes in the treatment group
- Conditional on margins, T has hypergeometric distribution under H_0
- The test may be quite conservative (due to discreteness)

Example Data

	Trt	**Ctrl**	Sum
Succ.	**3**	1	4
Fail	2	4	6
Sum	5	5	

Observed T = 3

Null distribution of T

k	**P(T=k)**
0	0.0238
1	0.2381
2	0.4762
3	0.2381
4	0.0238

Rejection region (k = 4)

p-value = 0.2619 (k = 3, 4)

Multiple Fisher Tests

- Consider two endpoints, A and B
- Four possible outcomes: No success, Success A only, Success B only, Success for A and B
- Global null-hypothesis: $H_0 : OR_A \leq 1$ and $OR_B \leq 1$
- Simple approach: Two marginal Fisher tests at level $\alpha/2$
- May be very conservative

Example Data

	Trt	**Ctrl**	Sum
AB	8	3	9
A	3	2	5
B	2	3	5
Fail	2	7	11
Sum	15	15	

Marginal Table EP A

	Trt	**Ctrl**	Sum
Succ.	**11**	5	16
Fail	4	10	14
Sum	15	15	

T_A = 11

p = 0.033 > 0.0125

Marginal Table EP B

	Trt	**Ctrl**	Sum
Succ.	**10**	6	16
Fail	5	9	14
Sum	15	15	

T_B = 10

p = 0.1362 > 0.0125

Idea: Use conditional joint distribution of the test statistics

	Trt	**Ctrl**	Sum
AB	x_1	y_1	a
A	x_2	y_2	b
B	x_3	y_3	c
Fail	x_4	y_4	d
Sum	n	m	N

- Fixed sample size per group n, m
- $X \sim Mult(n, p_{X_1}, p_{X_2}, p_{X_3}, p_{X_4})$,
- $Y \sim Mult(m, p_{Y_1}, p_{Y_2}, p_{Y_3}, p_{Y_4})$
- X, Y independent.
- $P(X_1 = x_1, X_2 = x_2, X_3 = x_3, X_4 = x_4 |$ $X_1 + Y_1 = a, X_2 + Y_2 = b, X_3 + Y_3 = c, X_4 + Y_4 = d) \propto$ $\propto \prod_{i=1}^{4} \frac{1}{x_i! y_i!} \left(\frac{p_{X_i}}{p_{Y_i}} \right)^{x_i}$
- Use this to find conditional joint distribution of the test statistics $T_A = x_1 + x_2$, $T_B = x_1 + x_3$
- Consider the point null-hypothesis $H_0 : p_{x_i} = p_{y_i}, \forall i = 1, ..., 4$
- Or any alternative specified through $p_{X_i}/p_{Y_i}, i = 1, ..., 4$

Conditional joint distribution under H_0

Joint distribution of T_A and T_B

Probabilities in %, rounded to 1 decimal digit. Entries "0" are small but positive.

T_B \ T_A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
	0	**0**	**0**	**0.4**	**2.8**	**10.3**	**22.1**	**28.5**	**22.1**	**10.3**	**2.8**	**0.4**	**0**	**0**	**0**	**100**
15										**0**	**0**					**0**
14									**0**	**0**	**0**	**0**				**0**
13								**0**	**0**	**0**	**0**	**0**	**0**			**0**
12							**0**	**0.1**	**0.1**	**0.1**	**0.1**	**0**	**0**	**0**		**0.4**
11						**0**	**0.3**	**0.7**	**0.9**	**0.6**	**0.2**	**0.1**	**0**	**0**	**0**	**2.8**
10					**0**	**0.4**	**1.5**	**2.8**	**3**	**1.8**	**0.6**	**0.1**	**0**	**0**	**0**	**10.3**
9				**0**	**0.3**	**1.5**	**4.1**	**6.4**	**5.8**	**3**	**0.9**	**0.1**	**0**	**0**		**22.1**
8			**0**	**0.1**	**0.7**	**2.8**	**6.4**	**8.4**	**6.4**	**2.8**	**0.7**	**0.1**	**0**			**28.5**
7		**0**	**0**	**0.1**	**0.9**	**3**	**5.8**	**6.4**	**4.1**	**1.5**	**0.3**	**0**				**22.1**
6	**0**	**0**	**0**	**0.1**	**0.6**	**1.8**	**3**	**2.8**	**1.5**	**0.4**	**0**					**10.3**
5	**0**	**0**	**0**	**0.1**	**0.2**	**0.6**	**0.9**	**0.7**	**0.3**	**0**						**2.8**
4		**0**	**0**	**0**	**0.1**	**0.1**	**0.1**	**0.1**	**0**							**0.4**
3			**0**	**0**	**0**	**0**	**0**	**0**								**0**
2				**0**	**0**	**0**	**0**									**0**
1					**0**	**0**										**0**

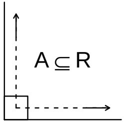
Under alternative of specific success probabilities

Succ. probs. for EP A and B: Treatment 0.7, 0.4, control 0.25, 0.25, correlation between endpoints 0.5

T_B \ T_A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
	0	**0**	**0**	**0**	**0**	**0**	**0.1**	**0.6**	**3.6**	**13**	**27.4**	**31.8**	**18.6**	**4.6**	**0.3**	**100**
15										**0**	**0**					**0**
14									**0**	**0**	**0**	**0**				**0**
13								**0**	**0**	**0.1**	**0.3**	**0.4**	**0.2**			**0.9**
12							**0**	**0**	**0.1**	**0.6**	**1.7**	**2.4**	**1.5**	**0.3**		**6.8**
11						**0**	**0**	**0**	**0.4**	**2**	**5.2**	**7.2**	**5**	**1.5**	**0.1**	**21.5**
10					**0**	**0**	**0**	**0.1**	**0.9**	**3.6**	**8.4**	**10.6**	**6.9**	**2**	**0.2**	**32.7**
9				**0**	**0**	**0**	**0**	**0.2**	**1.1**	**3.7**	**7.4**	**8**	**4.1**	**0.8**		**25.2**
8			**0**	**0**	**0**	**0**	**0**	**0.1**	**0.7**	**2.2**	**3.5**	**2.8**	**0.9**			**10.3**
7		**0**	**0**	**0**	**0**	**0**	**0**	**0.1**	**0.3**	**0.7**	**0.8**	**0.4**				**2.3**
6	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0.1**	**0.1**	**0.1**					**0.3**
5	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**						**0**
4		**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**							**0**
3			**0**	**0**	**0**	**0**	**0**	**0**								**0**
2				**0**	**0**	**0**	**0**									**0**
1					**0**	**0**										**0**

Definition of valid rejection regions

- A valid rejection region $R \in \mathbb{N} \times \mathbb{N}$ must satisfy
  - (i) Under H_0, $P((T_A, T_B) \in R) \leq \alpha$
  - (ii) If $(t_1, t_2) \in R$ then $A = \{(s_1, s_2) \in \mathbb{N} \times \mathbb{N} : s_1 \geq t_1, 2_2 \geq t_2\} \subseteq R$



- Condition (i) ensures level α control.
- Condition (ii) prevents implausible results and ensures larger power for larger effects.
- Reject H_0 if $(T_A, T_B) \in R$

Bonferroni rejection region for the example

Bonferroni rejection region

T_B																
	0	0	0.03	0.43	2.82	10.34	22.15	28.48	22.15	10.34	2.82	0.43	0.03	0	0	100
15										0	0					0
14									0	0	0	0				0
13								0	0.01	0.01	0.01	0	0			0.03
12							0.02	0.08	0.14	0.12	0.05	0.01	0	0		0.43
11						0.04	0.26	0.69	0.9	0.63	0.24	0.05	0.01	0	0	2.82
10					0.04	0.38	1.46	2.83	3.03	1.84	0.63	0.12	0.01	0	0	10.34
9				0.02	0.26	1.46	4.11	6.43	5.79	3.03	0.9	0.14	0.01	0		22.15
8			0	0.08	0.69	2.83	6.43	8.41	6.43	2.83	0.69	0.08	0			28.48
7		0	0.01	0.14	0.9	3.03	5.79	6.43	4.11	1.46	0.26	0.02				22.15
6	0	0	0.01	0.12	0.63	1.84	3.03	2.83	1.46	0.38	0.04					10.34
5	0	0	0.01	0.05	0.24	0.63	0.9	0.69	0.26	0.04						2.82
4		0	0	0.01	0.05	0.12	0.14	0.08	0.02							0.43
3			0	0	0.01	0.01	0.01	0								0.03
2				0	0	0	0									0
1					0	0										0
T_A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	

Alpha = 2.5 %; Level = 0.909 %; Area = 34

Aims: Use joint distribution to ...

- Find rejection region with maximal power for a specified alternative
- Also find rejection regions with
  - Maximal alpha exploitation
  - Maximal area
- Study special cases of exact p-value combination tests
  - $C = min(p_1, p_2)$
  - $C = p_1 * p_2$
  - Reject H_0 if c greater or equal the α quantile of the exact conditional null-distribution of C.
- Calculate power for Bonferroni related tests
  - Bonferroni
  - Tarone test: Test at level α/m. m is the number of tests that can become significant at level α/m. Choose m as large as possible. May add unused alpha to some individual tests.

Optimal rejection regions

An optimal rejection region R can be found as solution of

- $P((T_A, T_B) \in R) \rightarrow max$, under a specified alternative (maximal power)
- subject to conditions (i) and (ii)

Similarly we can optimize either

- $P((T_A, T_B) \in R) \rightarrow max$, under H_0 (maximal α)

or

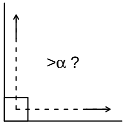
- $|R| \rightarrow max$ (maximal area)

This is a binary optimization problem, since each point (t_1, t_2) is either element of R or not element of R.

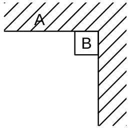
Optimization preprocessing

Size of search space is of order 2^n, preprocessing useful

1. Exclude points $(t_1, t_2) : P\{(s_1, s_2) : s_1 \geq t_1, s_2 \geq t_2\} > \alpha$



2. For all points $B = (t_1, t_2)$ get largest region A meeting condition (ii) without that point



If $P(A) + P(B) \leq \alpha$, $B \in R$ for sure

Preprocessing result example

Preprocessing

Excluded: Black, Remaining search space: Red, Included: Green

T_B \ T_A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
	0	0	0	0.4	2.8	10.3	22.1	28.5	22.1	10.3	2.8	0.4	0	0	0	100
15										0	0					0
14									0	0	0	0				0
13								0	0	0	0	0	0			0
12							0	0.1	0.1	0.1	0.1	0	0	0		0.4
11						0	0.3	0.7	0.9	0.6	0.2	0.1	0	0	0	2.8
10					0	0.4	1.5	2.8	3	1.8	0.6	0.1	0	0	0	10.3
9				0	0.3	1.5	4.1	6.4	5.8	3	0.9	0.1	0	0		22.1
8			0	0.1	0.7	2.8	6.4	8.4	6.4	2.8	0.7	0.1	0			28.5
7		0	0	0.1	0.9	3	5.8	6.4	4.1	1.5	0.3	0				22.1
6	0	0	0	0.1	0.6	1.8	3	2.8	1.5	0.4	0					10.3
5	0	0	0	0.1	0.2	0.6	0.9	0.7	0.3	0						2.8
4		0	0	0	0.1	0.1	0.1	0.1	0							0.4
3			0	0	0	0	0	0								0
2				0	0	0	0									0
1					0	0										0

Alpha = 2.5 %; Level = 3.591 %; Area = 16

Actual opimization as binary linear program

- Solution is of type $x = (0, 1, 0, \dots)$
- Objective function can be power, alpha, area
- γ_i contribution of element x_i to the objective function
- Find R, such that $\gamma^T x \rightarrow max$, s.t. conditions (i) and (ii)
- Condition (ii) imposes $x_i \leq x_j$ for certain pairs i, j
- Constraint matrix for linear program:

$$\begin{pmatrix} 1 & 1 & 1 & 1 & \dots & 1 \\ 1 & -1 & 0 & 0 & \dots & 0 \\ 1 & 0 & -1 & 0 & \dots & 0 \\ 0 & 1 & 0 & -1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} x \leq \begin{pmatrix} \alpha \\ 0 \\ 0 \\ 0 \\ \vdots \end{pmatrix}$$

- Use LP solver (branch and bound algorithm)

Result: Maximal power rejection region

Region with maximal power for specified alternative

Succ. probs. EP A, EP B: Trtm. 0.7, 0.4, Contr. 0.25, 0.25. Common correlation 0.5

T_B \ T_A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
	0	**0**	**0.03**	**0.43**	**2.82**	**10.34**	**22.15**	**28.48**	**22.15**	**10.34**	**2.82**	**0.43**	**0.03**	**0**	**0**	**100**
15										0	0					0
14									0	0	0	0				0
13								0	0.01	0.01	0.01	0	0			0.03
12							0.02	0.08	0.14	0.12	0.05	0.01	0	0		0.43
11						0.04	0.26	0.69	0.9	0.63	0.24	0.05	0.01	0	0	2.82
10					0.04	0.38	1.46	2.83	3.03	1.84	0.63	0.12	0.01	0	0	10.34
9				0.02	0.26	1.46	4.11	6.43	5.79	3.03	0.9	0.14	0.01	0		22.15
8			0	0.08	0.69	2.83	6.43	8.41	6.43	2.83	0.69	0.08	0			28.48
7		0	0.01	0.14	0.9	3.03	5.79	6.43	4.11	1.46	0.26	0.02				22.15
6	0	0	0.01	0.12	0.63	1.84	3.03	2.83	1.46	0.38	0.04					10.34
5	0	0	0.01	0.05	0.24	0.63	0.9	0.69	0.26	0.04						2.82
4		0	0	0.01	0.05	0.12	0.14	0.08	0.02							0.43
3			0	0	0.01	0.01	0.01	0								0.03
2				0	0	0	0									0
1					0	0										0

Alpha = 2.5 %; Level = 2.442 %; Area = 34

Maximal alpha rejection region

Rejection region with maximal alpha (red area)

T_B \ T_A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
	0	0	0.03	0.43	2.82	10.34	22.15	28.48	22.15	10.34	2.82	0.43	0.03	0	0	100
15										0	0					0
14									0	0	0	0				0
13								0	0.01	0.01	0.01	0	0			0.03
12							0.02	0.08	0.14	0.12	0.05	0.01	0	0		0.43
11						0.04	0.26	0.69	0.9	0.63	0.24	0.05	0.01	0	0	2.82
10					0.04	0.38	1.46	2.83	3.03	1.84	0.63	0.12	0.01	0	0	10.34
9				0.02	0.26	1.46	4.11	6.43	5.79	3.03	0.9	0.14	0.01	0		22.15
8			0	0.08	0.69	2.83	6.43	8.41	6.43	2.83	0.69	0.08	0			28.48
7		0	0.01	0.14	0.9	3.03	5.79	6.43	4.11	1.46	0.26	0.02				22.15
6	0	0	0.01	0.12	0.63	1.84	3.03	2.83	1.46	0.38	0.04					10.34
5	0	0	0.01	0.05	0.24	0.63	0.9	0.69	0.26	0.04						2.82
4		0	0	0.01	0.05	0.12	0.14	0.08	0.02							0.43
3			0	0	0.01	0.01	0.01	0								0.03
2				0	0	0	0									0
1					0	0										0

Alpha = 2.5 %; Level = 2.486 %; Area = 33

Maximal area rejection region

Region with maximal area

T_B \ T_A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
	0	0	0.03	0.43	2.82	10.34	22.15	28.48	22.15	10.34	2.82	0.43	0.03	0	0	100
15										0	0					0
14									0	0	0	0				0
13								0	0.01	0.01	0.01	0	0			0.03
12							0.02	0.08	0.14	0.12	0.05	0.01	0	0		0.43
11						0.04	0.26	0.69	0.9	0.63	0.24	0.05	0.01	0	0	2.82
10					0.04	0.38	1.46	2.83	3.03	1.84	0.63	0.12	0.01	0	0	10.34
9				0.02	0.26	1.46	4.11	6.43	5.79	3.03	0.9	0.14	0.01	0		22.15
8			0	0.08	0.69	2.83	6.43	8.41	6.43	2.83	0.69	0.08	0			28.48
7		0	0.01	0.14	0.9	3.03	5.79	6.43	4.11	1.46	0.26	0.02				22.15
6	0	0	0.01	0.12	0.63	1.84	3.03	2.83	1.46	0.38	0.04					10.34
5	0	0	0.01	0.05	0.24	0.63	0.9	0.69	0.26	0.04						2.82
4		0	0	0.01	0.05	0.12	0.14	0.08	0.02							0.43
3			0	0	0.01	0.01	0.01	0								0.03
2				0	0	0	0									0
1					0	0										0

Alpha = 2.5 %; Level = 2.42 %; Area = 37

$C = p_1 p_2$ rejection region

C=p_1p_2 rejection region

T_B \ T_A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
	0	0	0.03	0.43	2.82	10.34	22.15	28.48	22.15	10.34	2.82	0.43	0.03	0	0	100
15										0	0					0
14									0	0	0	0				0
13								0	0.01	0.01	0.01	0	0			0.03
12							0.02	0.08	0.14	0.12	0.05	0.01	0	0		0.43
11						0.04	0.26	0.69	0.9	0.63	0.24	0.05	0.01	0	0	2.82
10					0.04	0.38	1.46	2.83	3.03	1.84	0.63	0.12	0.01	0	0	10.34
9				0.02	0.26	1.46	4.11	6.43	5.79	3.03	0.9	0.14	0.01	0		22.15
8			0	0.08	0.69	2.83	6.43	8.41	6.43	2.83	0.69	0.08	0			28.48
7		0	0.01	0.14	0.9	3.03	5.79	6.43	4.11	1.46	0.26	0.02				22.15
6	0	0	0.01	0.12	0.63	1.84	3.03	2.83	1.46	0.38	0.04					10.34
5	0	0	0.01	0.05	0.24	0.63	0.9	0.69	0.26	0.04						2.82
4		0	0	0.01	0.05	0.12	0.14	0.08	0.02							0.43
3			0	0	0.01	0.01	0.01	0								0.03
2				0	0	0	0									0
1					0	0										0

Alpha = 2.5 %; Level = 2.42 %; Area = 37

$C = min(p_1, p_2)$ rejection region

C=min(p_1,p_2) rejection region

T_B \ T_A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
	0	0	0.03	0.43	2.82	10.34	22.15	28.48	22.15	10.34	2.82	0.43	0.03	0	0	100
15										0	0					0
14									0	0	0	0				0
13								0	0.01	0.01	0.01	0	0			0.03
12							0.02	0.08	0.14	0.12	0.05	0.01	0	0		0.43
11						0.04	0.26	0.69	0.9	0.63	0.24	0.05	0.01	0	0	2.82
10					0.04	0.38	1.46	2.83	3.03	1.84	0.63	0.12	0.01	0	0	10.34
9				0.02	0.26	1.46	4.11	6.43	5.79	3.03	0.9	0.14	0.01	0		22.15
8			0	0.08	0.69	2.83	6.43	8.41	6.43	2.83	0.69	0.08	0			28.48
7		0	0.01	0.14	0.9	3.03	5.79	6.43	4.11	1.46	0.26	0.02				22.15
6	0	0	0.01	0.12	0.63	1.84	3.03	2.83	1.46	0.38	0.04					10.34
5	0	0	0.01	0.05	0.24	0.63	0.9	0.69	0.26	0.04						2.82
4		0	0	0.01	0.05	0.12	0.14	0.08	0.02							0.43
3			0	0	0.01	0.01	0.01	0								0.03
2				0	0	0	0									0
1					0	0										0

Alpha = 2.5 %; Level = 0.909 %; Area = 34

Tarone rejection region

Tarone rejection region

T_B \ T_A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
	0	0	0.03	0.43	2.82	10.34	22.15	28.48	22.15	10.34	2.82	0.43	0.03	0	0	100
15										0	0					0
14									0	0	0	0				0
13								0	0.01	0.01	0.01	0	0			0.03
12							0.02	0.08	0.14	0.12	0.05	0.01	0	0		0.43
11						0.04	0.26	0.69	0.9	0.63	0.24	0.05	0.01	0	0	2.82
10					0.04	0.38	1.46	2.83	3.03	1.84	0.63	0.12	0.01	0	0	10.34
9				0.02	0.26	1.46	4.11	6.43	5.79	3.03	0.9	0.14	0.01	0		22.15
8			0	0.08	0.69	2.83	6.43	8.41	6.43	2.83	0.69	0.08	0			28.48
7		0	0.01	0.14	0.9	3.03	5.79	6.43	4.11	1.46	0.26	0.02				22.15
6	0	0	0.01	0.12	0.63	1.84	3.03	2.83	1.46	0.38	0.04					10.34
5	0	0	0.01	0.05	0.24	0.63	0.9	0.69	0.26	0.04						2.82
4		0	0	0.01	0.05	0.12	0.14	0.08	0.02							0.43
3			0	0	0.01	0.01	0.01	0								0.03
2				0	0	0	0									0
1					0	0										0

Alpha = 2.5 %; Level = 0.909 %; Area = 34

Power of Bonferroni rejection region = 58%

Power for Bonferroni region

Succ. probs. EP A, EP B: Trtm. 0.7, 0.4, Contr. 0.25, 0.25. Common correlation 0.5

T_B \ T_A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
	0	0	0	0	0	0	0.06	0.59	3.59	13.03	27.41	31.79	18.57	4.61	0.33	100
15										0	0					0
14									0	0.01	0.02	0.02				0.05
13								0	0.01	0.1	0.29	0.37	0.16			0.93
12							0	0.01	0.11	0.61	1.72	2.42	1.55	0.34		6.76
11						0	0	0.05	0.43	2.01	5.19	7.18	4.98	1.51	0.14	21.5
10					0	0	0.01	0.12	0.89	3.63	8.38	10.62	6.89	1.98	0.19	32.71
9				0	0	0	0.02	0.18	1.06	3.71	7.4	7.96	4.12	0.79		25.23
8			0	0	0	0	0.02	0.14	0.73	2.16	3.51	2.84	0.88			10.27
7		0	0	0	0	0	0.01	0.07	0.29	0.69	0.82	0.38				2.26
6	0	0	0	0	0	0	0	0.02	0.06	0.11	0.07					0.27
5	0	0	0	0	0	0	0	0	0.01	0.01						0.02
4		0	0	0	0	0	0	0	0							0
3			0	0	0	0	0	0								0
2				0	0	0	0									0
1					0	0										0

Alpha = 2.5 %; Level = 58.191 %; Area = 34

Power for maximal alpha region = 63%

Power for maximal alpha region

Succ. probs. EP A, EP B: Trtm. 0.7, 0.4, Contr. 0.25, 0.25. Common correlation 0.5

T_B \ T_A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
	0	0	0	0	0	0	0.06	0.59	3.59	13.03	27.41	31.79	18.57	4.61	0.33	100
15										0	0					0
14									0	0.01	0.02	0.02				0.05
13								0	0.01	0.1	0.29	0.37	0.16			0.93
12							0	0.01	0.11	0.61	1.72	2.42	1.55	0.34		6.76
11						0	0	0.05	0.43	2.01	5.19	7.18	4.98	1.51	0.14	21.5
10					0	0	0.01	0.12	0.89	3.63	8.38	10.62	6.89	1.98	0.19	32.71
9				0	0	0	0.02	0.18	1.06	3.71	7.4	7.96	4.12	0.79		25.23
8			0	0	0	0	0.02	0.14	0.73	2.16	3.51	2.84	0.88			10.27
7		0	0	0	0	0	0.01	0.07	0.29	0.69	0.82	0.38				2.26
6	0	0	0	0	0	0	0	0.02	0.06	0.11	0.07					0.27
5	0	0	0	0	0	0	0	0	0.01	0.01						0.02
4		0	0	0	0	0	0	0	0							0
3			0	0	0	0	0	0								0
2				0	0	0	0									0
1					0	0										0

Alpha = 2.5 %; Level = 62.597 %; Area = 33

Power for maximal area region = 73%

Power for maximal area region

Succ. probs. EP A, EP B: Trtm. 0.7, 0.4, Contr. 0.25, 0.25. Common correlation 0.5

T_B \ T_A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
	0	**0**	**0**	**0**	**0**	**0**	**0.06**	**0.59**	**3.59**	**13.03**	**27.41**	**31.79**	**18.57**	**4.61**	**0.33**	**100**
15										0	0					0
14									0	0.01	0.02	0.02				0.05
13								0	0.01	0.1	0.29	0.37	0.16			0.93
12							0	0.01	0.11	0.61	1.72	2.42	1.55	0.34		6.76
11						0	0	0.05	0.43	2.01	5.19	7.18	4.98	1.51	0.14	21.5
10					0	0	0.01	0.12	0.89	3.63	8.38	10.62	6.89	1.98	0.19	32.71
9				0	0	0	0.02	0.18	1.06	3.71	7.4	7.96	4.12	0.79		25.23
8			0	0	0	0	0.02	0.14	0.73	2.16	3.51	2.84	0.88			10.27
7		0	0	0	0	0	0.01	0.07	0.29	0.69	0.82	0.38				2.26
6	0	0	0	0	0	0	0	0.02	0.06	0.11	0.07					0.27
5	0	0	0	0	0	0	0	0	0.01	0.01						0.02
4		0	0	0	0	0	0	0	0							0
3			0	0	0	0	0	0								0
2				0	0	0	0									0
1					0	0										0

Alpha = 2.5 %; Level = 73.775 %; Area = 37

Maximal power = 79%

Power for maximal power region

Succ. probs. EP A, EP B: Trtm. 0.7, 0.4, Contr. 0.25, 0.25. Common correlation 0.5

T_B \ T_A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
	0	0	0	0	0	0	0.06	0.59	3.59	13.03	27.41	31.79	18.57	4.61	0.33	100
15										0	0					0
14									0	0.01	0.02	0.02				0.05
13								0	0.01	0.1	0.29	0.37	0.16			0.93
12							0	0.01	0.11	0.61	1.72	2.42	1.55	0.34		6.76
11						0	0	0.05	0.43	2.01	5.19	7.18	4.98	1.51	0.14	21.5
10					0	0	0.01	0.12	0.89	3.63	8.38	10.62	6.89	1.98	0.19	32.71
9				0	0	0	0.02	0.18	1.06	3.71	7.4	7.96	4.12	0.79		25.23
8			0	0	0	0	0.02	0.14	0.73	2.16	3.51	2.84	0.88			10.27
7		0	0	0	0	0	0.01	0.07	0.29	0.69	0.82	0.38				2.26
6	0	0	0	0	0	0	0	0.02	0.06	0.11	0.07					0.27
5	0	0	0	0	0	0	0	0	0.01	0.01						0.02
4		0	0	0	0	0	0	0	0							0
3			0	0	0	0	0	0								0
2				0	0	0	0									0
1					0	0										0

Alpha = 2.5 %; Level = 79.046 %; Area = 34

Exact power (%) for different alternatives

n=15 per group

	Succ. prob. treat.		Optimize			Combine p-values		Bonferroni-related		Simplifications		
Corr.	**EP A**	**EP B**	**Power**	**Area**	**Alpha**	**p1*p2**	**minP**	**Tarone**	**Bonf.**	**Collapse**	**Only A**	**Only B**
0	0.25	0.25	2.4	2.2	2.4	2.0	1.3	1.4	0.8	0.8	0.9	0.9
0	0.75	0.25	76.8	67.8	52.5	69.9	67.5	67.7	55.0	39.7	67.5	0.9
0	0.75	0.5	87.8	81.6	77.0	84.0	70.8	72.1	59.8	58.3	67.5	15.7
0	0.75	0.75	97.7	95.7	95.4	96.8	87.4	88.2	79.6	79.6	67.5	67.5
0.5	0.25	0.25	2.4	2.2	2.4	2.0	1.3	1.4	0.8	0.8	0.9	0.9
0.5	0.75	0.25	75.0	66.6	51.3	67.4	67.5	67.7	55.0	39.0	67.5	0.9
0.5	0.75	0.5	85.3	79.1	74.6	81.1	70.3	71.3	59.1	54.8	67.5	15.7
0.5	0.75	0.75	96.4	94.3	94.2	95.5	86.5	86.9	78.1	76.7	67.5	67.5

Success probability control = 0.25 for both EP

- For these scenarios, maximal power test is 0.9 to 7.3 % points better than next best test
- Maximizing alpha is not sole the key to good power
- $C = p_1 p_2$ and maximal area test similar
- Tarone and minP very similar
- Simplifications have least power

Non-inferiority - superiority tests

What if we want to show non-inferiority for both EP and superiority in the sense of the global null-hypothesis for some EP?

- Find optimal rejection region for global superiority test within the rejection region of the non-inferiority test.
- Problem: At the moment the optimal test is designed for a point null-hypothesis
- Type-I-error rate control for more general null-hypothesis $H_0 : OR_1 \leq 1$ AND $OR_2 \leq 1$?

Discussion

- Exact conditional tests useful for small sample inference
- Example of two binary endpoints
- Can get optimal rejection region for a specified alternative
  - Generalizations to more endpoints no problem in principle
  - But optimization run-time may be non-polynomial
- Level of significance close to nominal level
  - Not sole determinant of good power
  - Behavior under other than point null-hypothesis remains to be analyzed
- Exact joint distribution can be used to study other tests
  - Maxmial power test can be used as benchmark
- Further restictions can be implemented, e.g. as suggested for non-inferiority - superiority test

Literature

Westfall, P. H., Young, S. S. (1989). P value adjustments for multiple tests in multivariate binomial models. Journal of the American Statistical Association, 84(407), 780-786

Tarone, R. E. (1990). A modified Bonferroni method for discrete data. Biometrics, 515-522.

Agresti, A. (1992). A survey of exact inference for contingency tables. Statistical Science, 131-153.

Agresti, A. (2002). Categorical data analysis. A John Wiley and Sons, Inc. Publication, Hoboken, New Jersey, USA.

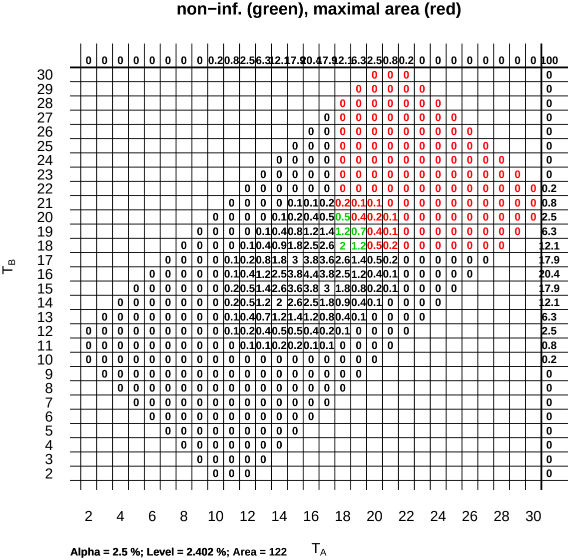
Wellek, S. (2005). Statistical Methods for the Analysis of Two Arm Non Inferiority Trials with Binary Outcomes. Biometrical journal, 47(1), 48-61.

Exact power (%) for different alternatives

n=10 per group

	Succ. prob. treat.		Optimize			Combine p-values		Bonferroni-related		Simplifications		
Corr.	**EP A**	**EP B**	**Power**	**Area**	**Alpha**	**p1*p2**	**minP**	**Tarone**	**Bonf.**	**Collapse**	**Only A**	**Only B**
0	0.25	0.25	2.0	1.9	2.0	1.7	0.9	0.9	0.4	0.6	0.5	0.5
0	0.75	0.25	54.1	47.2	41.3	48.4	42.2	42.3	41.6	22.0	42.1	0.5
0	0.75	0.5	69.1	62.0	61.0	63.0	46.6	46.6	45.7	34.2	42.1	8.6
0	0.75	0.75	87.5	83.4	84.1	84.5	66.3	66.3	65.8	51.6	42.1	42.1
0.5	0.25	0.25	2.0	1.9	2.0	1.7	0.9	0.9	0.4	0.6	0.5	0.5
0.5	0.75	0.25	51.8	45.6	39.4	46.8	42.2	42.3	41.6	21.7	42.1	0.5
0.5	0.75	0.5	65.5	59.1	57.8	59.7	46.1	46.1	45.1	31.9	42.1	8.6
0.5	0.75	0.75	84.1	80.7	81.4	81.2	64.8	64.8	64.3	49.4	42.1	42.1

Non-inferiority - superiority maximal area



Non-inferiority - superiority $C = p_1 p_2$