Kompetenzzentrum für Klinische Studien Bremen

The Populationwise Error Rate - A More Liberal Error Rate for Multiplicity Adjustment in Enrichment Designs

Kornelius Rohmeyer, Werner Brannath, Sarah Nanninga

Adaptive Designs and Multiple Testing Procedures Workshop
University of Padua, Italy
April 28th-29th 2016

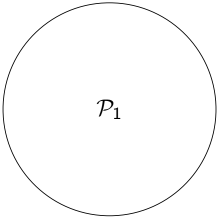
Overview

1 Setting & Definition

2 Closure principle for the PWER

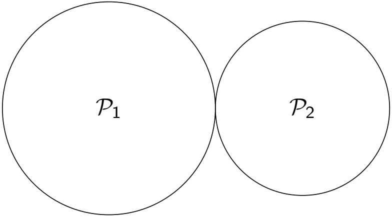
3 Estimation of population proportions π_i

Multiple Testing Error Correction



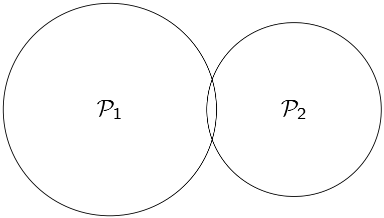
- One population with two hypotheses $\Rightarrow$ FWER

Multiple Testing Error Correction



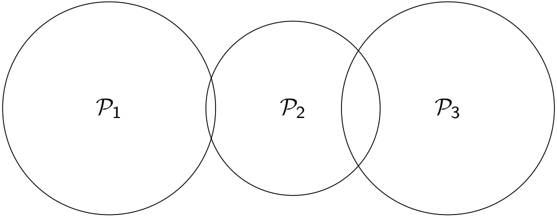
- One population with two hypotheses $\Rightarrow$ FWER
- Two disjunct populations with two hypotheses $\Rightarrow$ no correction

Multiple Testing Error Correction



- One population with two hypotheses $\Rightarrow$ FWER
- Two disjunct populations with two hypotheses $\Rightarrow$ no correction
- Two non-disjunct populations with two hypotheses $\Rightarrow$ FWER?

Multiple Testing Error Correction



- One population with two hypotheses $\Rightarrow$ FWER
- Two disjunct populations with two hypotheses $\Rightarrow$ no correction
- Two non-disjunct populations with two hypotheses $\Rightarrow$ FWER?
- Even if the FWER is correcting for some multiplicity no one is actually effected by?

PWER - Goal & Setting

Especially in small populations (for example defined by biomarkers) an error rate that is not as strict as the FWER, but continuously extends the spectrum from the FWER to the unadjusted case for disjunct populations may be of interest.

Our setting: $\mathcal{P}_1$, $\mathcal{P}_2$, ..., $\mathcal{P}_k$ different generally non-disjunct subpopulations with one hypothesis H_i for each $\mathcal{P}_i$, which is tested with a test φ_i.

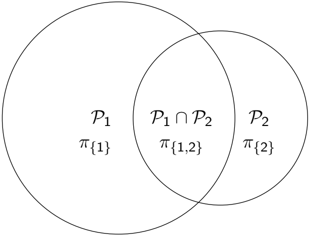
PWER - Populationwise Error Rate

For a random subject we regard the belonging to a certain subpopulation $\mathcal{P}_i$ as a random indicator variable $\Psi_i : \Omega \to \{0, 1\}$. We want to control the probability that a *random* person is effected by an erroneously rejected H_i. That is

$$P_\theta \left(\bigcup_{j \in I(\theta)} \{\varphi_j = 1\} \cap \{\Psi_j = 1\} \right)$$

with $I(\theta)$ the index set of the true null hypotheses under θ.

PWER - Example

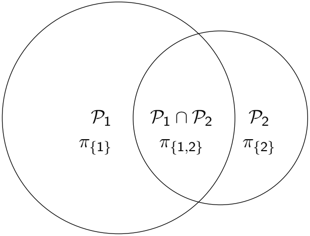


Let $\pi_{\{1\}}, \pi_{\{2\}}$ and $\pi_{\{1,2\}}$ be the proportions of *the partition* $\mathcal{P}_1 \setminus \mathcal{P}_2, \mathcal{P}_2 \setminus \mathcal{P}_1$ and $\mathcal{P}_1 \cap \mathcal{P}_2$ respectively.

Under global H_0:

$$\mathrm{PWER} = \pi_{\{1\}} \cdot P(\varphi_1 = 1) + \pi_{\{2\}} \cdot P(\varphi_2 = 1) + \pi_{\{1,2\}} \cdot P(\max_{i=1,2} \varphi_i = 1)$$

PWER - Example



Let $\pi_{\{1\}}, \pi_{\{2\}}$ and $\pi_{\{1,2\}}$ be the proportions of *the partition* $\mathcal{P}_1 \setminus \mathcal{P}_2, \mathcal{P}_2 \setminus \mathcal{P}_1$ and $\mathcal{P}_1 \cap \mathcal{P}_2$ respectively.

Under global H_0:

$$\text{PWER} = \pi_{\{1\}} \cdot P(\varphi_1 = 1) + \pi_{\{2\}} \cdot P(\varphi_2 = 1) + \pi_{\{1,2\}} \cdot P(\max_{i=1,2} \varphi_i = 1)$$

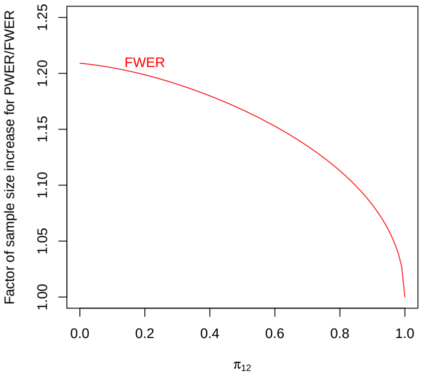
If the aim is to reject complete populations $\mathcal{P}_1$ or $\mathcal{P}_2$
$\Rightarrow$ equal critical values for test statistics for φ_1, φ_2 and $\max_{i=1,2} \varphi_i$.

PWER - How do power & sample size compare to the FWER?

Assumptions

- φ_1, φ_2 z-tests.
- $\mathrm{Cor}(\varphi_1, \varphi_2) = \frac{\pi_{\{1,2\}}}{\sqrt{\pi_1 \pi_2}}$
- Proportion $\pi_{\{1\}} : \pi_{\{2\}} = 1 : 1$

Graph: Depending on $\pi_{\{1,2\}}$ how does the sample size increase when the FWER/PWER is controlled but the same marginal power should be achieved as in the unadjusted case.

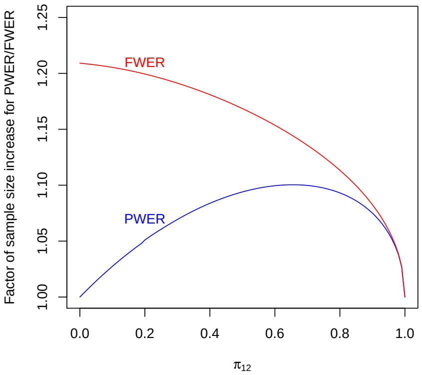


PWER - How do power & sample size compare to the FWER?

Assumptions

- φ_1, φ_2 z-tests.
- $\text{Cor}(\varphi_1, \varphi_2) = \frac{\pi_{\{1,2\}}}{\sqrt{\pi_1 \pi_2}}$
- Proportion
  $\pi_{\{1\}} : \pi_{\{2\}} = 1 : 1$

Graph: Depending on $\pi_{\{1,2\}}$ how does the sample size increase when the FWER/PWER is controlled but the same marginal power should be achieved as in the unadjusted case.



Closure principle for the PWER

Theorem (Variant I)

For each $J \subset I := \{1, \ldots, k\}$ *let*

$$(\varphi_j^J \ : \ j \in J)$$

be a family of tests.
If $(\varphi_j^J \ : \ j \in J)$ *controls the* PWER *weakly for* $H_j, j \in J$ *to level* α, *i.e.*

$$\sum_{\tilde{J} \subset J} \pi_{\tilde{J}} \cdot P_\theta \left(\{ \max_{j \in \tilde{J}} \varphi_j^J = 1 \} \right) \leq \alpha \quad \textit{for all} \quad \theta \in \cap_{j \in J} H_j,$$

then, with $\varphi_J := \max_{j \in J} \varphi_j^J$, *the tests*

$$\tilde{\varphi}_i := \min_{J:\, i \in J} \varphi_J, \quad i \in I$$

control the PWER *(strongly) for all* $\theta \in \Theta$ *at level* α.

Closure principle for the PWER

Theorem (Variant II)

Let $\varphi = (\varphi_J)_{J \subset I}$ *be an (arbitrary) family of tests for* $\mathcal{H} = \{H_J := \bigcap_{j \in J} H_j \mid J \subset I\}$, *where* φ_J *has a local level of* α_J.
If for all $\tilde{I} \subset I$

$$\sum_{J \subset \tilde{I}} \pi_J \min_{\tilde{I} \supset \tilde{J} \supset J} (\alpha_{\tilde{J}}) \leq \alpha \qquad (1)$$

then the tests

$$\tilde{\varphi}_i = \min_{J:\, i \in J} \varphi_J, \quad i \in I$$

control the PWER at level α.

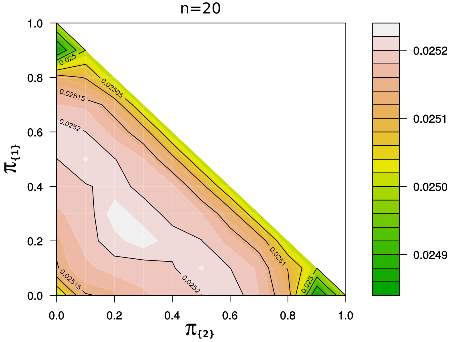
Unkown population proportions π_i

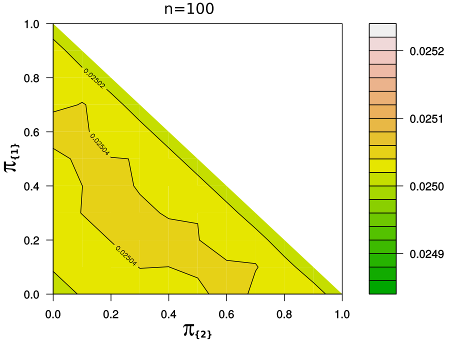
Problem: Normally we don't know the proportions π_i of the populations (and their intersections), but they are used in the PWER definition as weights for the different type I errors as well as they influence the correlation of the test statistics.
Possible approaches:

- Estimate π_i.
- Use of confidence regions and least favorable configurations in the confidence set.
- Bayesian approach for π_i.

Unkown population proportions π_i

When using the relative frequency $\hat{\pi}_i$ as if it is π_i, simulations show that the α-level is only increased slightly.





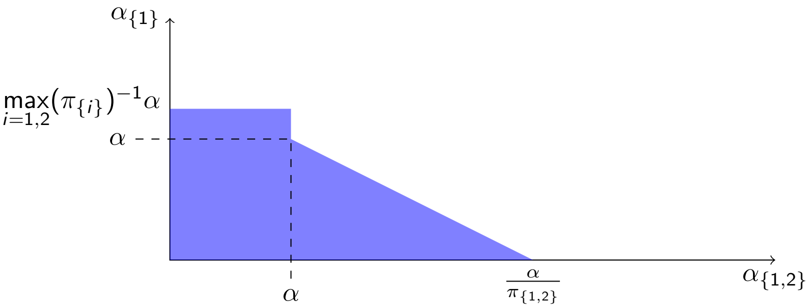
Outlook

- R package `pwer` in work.
- Group sequential and adaptive designs controlling the PWER.
  - Updated estimates of population proportions.
  - Stepwise increase of error rate stringency.

Thank you

Thank you for listening!

Admissible local α levels



That is, if we sacrifice some power for the elementary hypotheses, i.e. set $\alpha_{\{1\}} < \alpha$, we can choose $\alpha_{\{1,2\}}$ greater than α.