Multi-arm trials

Pavel Mozgunov

MRC Biostatistics Unit, University of Cambridge, UK,

22 October 2025

Multiple treatments

We have K experimental treatments $T_1, \ldots, T_K$ and control available.

Suppose that response to T_k are

$$Y_{ki} \sim \mathcal{N}(\mu_k, \sigma^2)$$

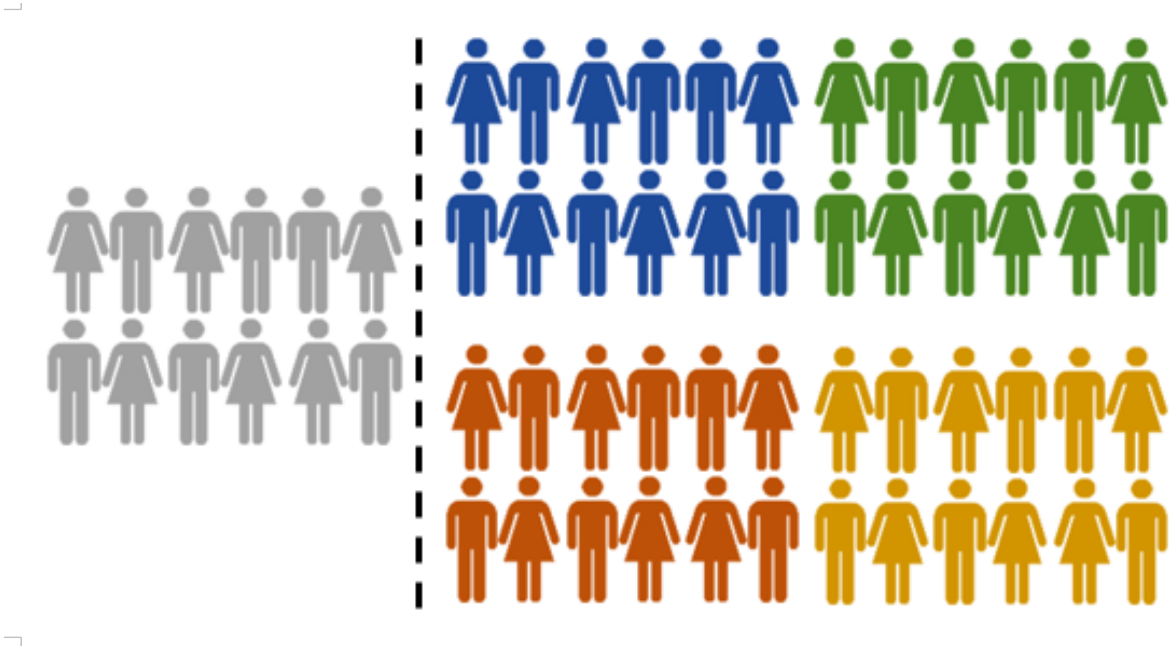
Interest is in testing K distinct null hypotheses

$$H_{0k} : \mu_0 \geq \mu_k \qquad H_{1k} : \mu_0 < \mu_k$$

for $k = 1, \ldots, K$ and where T_0 is the control treatment.

A simple multi-arm study

In a multi-arm study we only recruit from patients to the control treatment once, using a common control for the trial



Estimation

For each arm $k = 0, \ldots, K$ we find estimates $\bar{y}_k$ and corresponding Z-values for $k = 1, \ldots, K$

$$Z_k = \frac{1}{\sigma}\sqrt{\frac{n_0 n_k}{n_0 + n_k}}(\bar{y}_k - \bar{y}_0)$$

where n_k is the number of patients treated on T_k and

$$\boldsymbol{Z} \sim N(\boldsymbol{d}, \boldsymbol{V})$$

with

$$d_k = \frac{1}{\sigma}\sqrt{\frac{n_0 n_k}{n_0 + n_k}}(\mu_k - \mu_0); \quad v_{kk'} = \sqrt{\frac{n_k n_{k'}}{(n_0 + n_k)(n_0 + n_{k'})}},\ k \neq k'; \quad v_{kk} = 1$$

Selecting the best treatment

Select treatment T_{k*} where

$$Z_{k^*} = \max(Z_1, \ldots, Z_K)$$

provided that the null hypothesis H_{0k} is rejected

$$Z_{k^*} \geq c$$

Otherwise, make no selection

Specifying the accuracy

Under the global null:

$$\nu \quad = \quad P(\text{select a treatment} \mid \mu_0 = \mu_1 = \cdots = \mu_K)$$

Under the alternative:

$$1-\eta \quad = \quad P(\text{select } T_1 \mid \mu_1 = \mu_0 + \delta \text{ and } \mu_{k'} = \mu_0 + \delta_0,\ k' = 2, \ldots, K)$$

with $0 \leq \delta_0 < \delta$.

- δ is the interesting treatment effect
- δ_0 is the uninteresting treatment effect
- This is typically referred to as a *Least Favourable Configuration*

Set ν to be small and $1-\eta$ to be large

Familywise Error Rate (FWER)

Given multiple hypotheses the simply controlling the type I error rate is not sufficient, we require strong control of the FWER

Let $\boldsymbol{\mu} = (\mu_0, \mu_1, \ldots, \mu_K)$ then strong control is achieved when

$$\mathbb{P}(\text{Reject any combination of true null hypotheses}) \leq \nu \text{ for all } \boldsymbol{\mu}.$$

The Dunnett approach

Suppose we are testing H_{0,k^*}, let

$$Z_{k^*} = \text{argmax}_{k \in 1,\ldots,K} Z_k$$

choosing c such that when $\mu_k \geq \mu_0$, we have

$$\mathbb{P}(Z_{k^*} > c) \leq \nu$$

we reject H_{0k^*} when $Z_{k^*} > c$.

Syntax in R

R code

```
> #find the critical region for a test
> alpha<-0.05
> qnorm(1-alpha)
[1] 1.644854
```

R code

```
  > #find the critical region for correlated tests
> require(mvtnorm)
> corr<-matrix(c(1,0.5,0.5,1),ncol=2)
> corr
     [,1] [,2]
[1,]  1.0  0.5
[2,]  0.5  1.0
> qmvnorm((1-alpha),tail="lower.tail",
mean=c(0,0),corr=corr)$quantile
[1] 1.916399
```

Where simulation can help

- Assessment of power for individual hypotheses;
- Behaviour under different configurations (for the given sample size);
- Operating characteristics under non-normal endpoints;
- Accuracy of assumption of equal variances;
- ...

Your best friends for the rest of the day....

Simulating (random) outcomes:

R code

```
> rnorm(n=5,mean=0,sd=1)
[1] 0.8535184 0.1053162 2.4155607 1.9475439 0.4064240

> rmvnorm(n=3,mean=c(0,0),sigma=corr)
           [,1]       [,2]
[1,]  1.0394488  0.5164349
[2,]  0.3408723  0.8562483
[3,]  0.5389829 -0.1303348
```

"For" loops

R code

```
for (i in 1:nsims){
...}
```

Dom and Pavel