# Decision making in phase II/III trials using early endpoint data


## Nigel Stallard

Statistics & Epidemiology, Division of Health Sciences
Warwick Medical School, University of Warwick, UK
*n.stallard@warwick.ac.uk*

# Acknowledgments


## Collaborators

Sue Todd (Reading)

Tim Friede (UMG Göttingen)

Cornelia Kunz (Boeringher Ingleheim, formerly Warwick)

## Funding

UK Medical Research Council

# 1. Phase II/III trials - motivation

## 1.1 Setting

A trial in Alzheimer's disease

- Galantamine at 3 doses and placebo control

Primary endpoint

- ADAS-cog change over 12 weeks

Wilkinson D, Murray J. (2001) Galantamine: a randomized, double-blind, dose comparison in patients with Alzheimer's disease. *Int. J. Ger. Psych.*, **16**, 852-857.

# 1. Phase II/III trials - motivation

## 1.1 Setting

A trial in Alzheimer's disease
- Galantamine at 3 doses and placebo control

Primary endpoint
- ADAS-cog change over 12 weeks

Aims
- select most effective dose
- provide a valid comparison with control

Wilkinson D, Murray J. (2001) Galantamine: a randomized, double-blind, dose comparison in patients with Alzheimer's disease. *Int. J. Ger. Psych.*, **16**, 852-857.

# 1.2 Conventional approach

## Phase II trial

- exploratory trial compares three doses (with placebo)
- select best dose
- go on to phase III if sufficiently promising
- short-term endpoint: ADAS-cog change over 6 weeks

# 1.2 Conventional approach

## Phase II trial

- exploratory trial compares three doses (with placebo)
- select best dose
- go on to phase III if sufficiently promising
- short-term endpoint: ADAS-cog change over 6 weeks

## Phase III trial

- compares single selected dose with placebo
- long-term endpoint: ADAS-cog change over 12 weeks
- control error rates

# 1.3 Combining phases II and III

Trial is conducted in two stages

## Stage 1

- 3 doses + placebo
- short and long term endpoints
- conduct interim analysis to select most promising dose

# 1.3 Combining phases II and III

Trial is conducted in two stages

## Stage 1

- 3 doses + placebo
- short and long term endpoints
- conduct interim analysis to select most promising dose

## Stage 2

- selected dose + placebo
- long term endpoint

# 1.3 Combining phases II and III

Trial is conducted in two stages

## Stage 1

- 3 doses + placebo
- short and long term endpoints
- conduct interim analysis to select most promising dose

## Stage 2

- selected dose + placebo
- long term endpoint

## Final analysis

- use all long term endpoint data on selected dose
- control overall type I error rate providing valid comparison

# 2. Group-sequential phase III method

## 2.1 Notation and test statistics

Treatments: $T_0$ = control, $T_1$ = experimental

Data $\sim N(\mu_i, \sigma^2)$ for treatment $T_i, i = 0, 1$ (known $\sigma^2$)

$\theta = \mu_1 - \mu_0$ is a measure of superiority of $T_1$ over $T_0$

Jennison C, Turnbull BW. (2000) *Group sequential methods with applications to clinical trials.* Chapman and Hall: London.

# 2. Group-sequential phase III method

## 2.1 Notation and test statistics

Treatments: $T_0$ = control, $T_1$ = experimental

Data $\sim N(\mu_i, \sigma^2)$ for treatment $T_i, i = 0, 1$ (known $\sigma^2$)

$\theta = \mu_1 - \mu_0$ is a measure of superiority of $T_1$ over $T_0$

Stage $j$ $(j = 1, 2)$:

cumulative data from total of $n_j$ patients per treatment

obtain $\hat{\theta}_j$ and $var(\hat{\theta}_j) = 2\sigma^2/n_j$

Jennison C, Turnbull BW. (2000) *Group sequential methods with applications to clinical trials.* Chapman and Hall: London.

# 2. Group-sequential phase III method

## 2.1 Notation and test statistics

Treatments: $T_0$ = control, $T_1$ = experimental

Data $\sim N(\mu_i, \sigma^2)$ for treatment $T_i, i = 0, 1$ (known $\sigma^2$)

$\theta = \mu_1 - \mu_0$ is a measure of superiority of $T_1$ over $T_0$

Stage $j$ $(j = 1, 2)$:

cumulative data from total of $n_j$ patients per treatment

obtain $\hat{\theta}_j$ and $var(\hat{\theta}_j) = 2\sigma^2/n_j$

$S_j = \hat{\theta}_j / var(\hat{\theta}_j)$ is efficient score statistic

$I_j = 1/var(\hat{\theta}_j)$ is Fisher's information

Jennison C, Turnbull BW. (2000) *Group sequential methods with applications to clinical trials.* Chapman and Hall: London.

# 2.2 Stopping boundary

Test $H_0 : \theta = 0$

At first interim analysis

- $S_1 \geq u_1$: stop and reject $H_0$
- $S_1 \leq l_1$: stop do not reject $H_0$ (one-sided test)
- $l_1 < S_1 < u_1$: continue to next interim analysis

At second interim (final) analysis

- $S_2 \geq u_2$: stop and reject $H_0$
- $S_2 < l_2 = u_2$: stop do not reject $H_0$

# 2.3 Spending function

Stopping boundary determined by an $\alpha$-*spending function*

Specify $0 \leq \alpha^*(1) \leq \alpha^*(2) = \alpha$

Find $u_1, u_2$ such that for $j = 1, 2$

$$pr_{H_0}(S_1 \geq u_1) = \alpha^*(1) \tag{1}$$

$$pr_{H_0}(S_1 \geq u_1) + pr_{H_0}(S_2 \geq u_2, l_1 < S_1 < u_1) = \alpha^*(2) \tag{2}$$

# 2.3 Spending function

Stopping boundary determined by an $\alpha$-*spending function*

Specify $0 \leq \alpha^*(1) \leq \alpha^*(2) = \alpha$
Find $u_1, u_2$ such that for $j = 1, 2$

$$pr_{H_0}(S_1 \geq u_1) = \alpha^*(1) \quad (1)$$

$$pr_{H_0}(S_1 \geq u_1) + pr_{H_0}(S_2 \geq u_2, l_1 < S_1 < u_1) = \alpha^*(2) \quad (2)$$

($l_1$ specified by similar spending function or given in advance)

# 2.4 Stopping boundary calculations

Obtain boundary by considering density of $S_1$ and $S_2$

# 2.4 Stopping boundary calculations

Obtain boundary by considering density of $S_1$ and $S_2$

$$\begin{pmatrix} S_1 \\ S_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \theta I_1 \\ \theta I_2 \end{pmatrix}, \begin{pmatrix} I_1 & I_1 \\ I_1 & I_2 \end{pmatrix}\right)$$

# 2.4 Stopping boundary calculations

Obtain boundary by considering density of $S_1$ and $S_2$

$$\begin{pmatrix} S_1 \\ S_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \theta I_1 \\ \theta I_2 \end{pmatrix}, \begin{pmatrix} I_1 & I_1 \\ I_1 & I_2 \end{pmatrix}\right)$$

First analysis: under $H_0$, $S_1 \sim N(0, I_1)$

$$f_1(s_1) = \frac{1}{\sqrt{I_1}}\phi\left(\frac{s_1}{\sqrt{I_1}}\right)$$

Find $u_1$ to satisfy (1)

# 2.4 Stopping boundary calculations

Obtain boundary by considering density of $S_1$ and $S_2$

$$\begin{pmatrix} S_1 \\ S_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \theta I_1 \\ \theta I_2 \end{pmatrix}, \begin{pmatrix} I_1 & I_1 \\ I_1 & I_2 \end{pmatrix}\right)$$

First analysis: under $H_0$, $S_1 \sim N(0, I_1)$

$$f_1(s_1) = \frac{1}{\sqrt{I_1}}\phi\left(\frac{s_1}{\sqrt{I_1}}\right)$$

Find $u_1$ to satisfy (1)

Second analysis: $S_2$ is sum of $S_1 \in (l_1, u_1)$ and $N(0, I_2 - I_1)$

$$f_2(s_2) = \int_{l_1}^{u_1} \frac{1}{\sqrt{I_2 - I_1}}\phi\left(\frac{s_2 - s_1}{\sqrt{I_2 - I_1}}\right) f_1(s_1) ds_1$$

Find $u_2$ to satisfy (2)

# 3. Adaptive seamless phase II/III trial

## 3.1 Idea

Start with $k$ experimental treatments and control

At interim analysis:

- select treatment with largest estimated effect
- compare test statistic for this treatment with boundary
- stop if appropriate
- otherwise continue with this treatment + control

Stallard N, Todd S. (2003) Sequential designs for phase III clinical trials incorporating treatment selection. *Stat. Med.*, **22**, 689–703.

# 3. Adaptive seamless phase II/III trial

## 3.1 Idea

Start with $k$ experimental treatments and control

At interim analysis:

- select treatment with largest estimated effect
- compare test statistic for this treatment with boundary
- stop if appropriate
- otherwise continue with this treatment + control

Final analysis allows for interim analyses and selection (here with same endpoint at interim and final analyses)

Stallard N, Todd S. (2003) Sequential designs for phase III clinical trials incorporating treatment selection. *Stat. Med.*, **22**, 689–703.

# 3.2 Notation and test statistics

Treatments: $T_0$ = control, $T_1, \ldots, T_k$ = experimental
Normally distributed data with mean $\mu_i$ for treatment $T_i, i = 0, \ldots, k$ and known common variance

$\theta_i = \mu_i - \mu_0$ is a measure of superiority of $T_i$ over $T_0$
test global null hypothesis $H_0 : \theta_1 = \cdots = \theta_k = 0$

# 3.2 Notation and test statistics

Treatments: $T_0$ = control, $T_1, \ldots, T_k$ = experimental
Normally distributed data with mean $\mu_i$ for treatment $T_i, i = 0, \ldots, k$ and known common variance

$\theta_i = \mu_i - \mu_0$ is a measure of superiority of $T_i$ over $T_0$
test global null hypothesis $H_0 : \theta_1 = \cdots = \theta_k = 0$

Stage $j$: data from total of $n_j$ patients per treatment

- obtain $\hat{\theta}_{ij}$ and $var(\hat{\theta}_{ij}) = 2\sigma^2/n_j$
- $S_{ij} = \hat{\theta}_{ij}/var(\hat{\theta}_{ij})$ efficient score statistic for $\theta_i$
- $I_j = 1/var(\hat{\theta}_{ij})$ Fisher's information for $\theta_i$ (same for all $i$)

Consider density of $S_{ij}$, $i = 1, \ldots, k$, $j = 1, 2$ under $H_0$

$$\begin{pmatrix} S_{11} \\ \vdots \\ S_{k1} \end{pmatrix} \sim N\left( \begin{pmatrix} \theta_1 I_1 \\ \vdots \\ \theta_k I_1 \end{pmatrix}, \begin{pmatrix} I_1 & I_1/2 & \cdots & I_1/2 \\ I_1/2 & I_1 & \cdots & I_1/2 \\ \vdots & & \ddots & \vdots \\ I_1/2 & I_1/2 & \cdots & I_1 \end{pmatrix} \right)$$

Consider density of $S_{ij}$, $i = 1, \ldots, k$, $j = 1, 2$ under $H_0$

$$\begin{pmatrix} S_{11} \\ \vdots \\ S_{k1} \end{pmatrix} \sim N\left( \begin{pmatrix} \theta_1 I_1 \\ \vdots \\ \theta_k I_1 \end{pmatrix}, \begin{pmatrix} I_1 & I_1/2 & \cdots & I_1/2 \\ I_1/2 & I_1 & \cdots & I_1/2 \\ \vdots & & \ddots & \vdots \\ I_1/2 & I_1/2 & \cdots & I_1 \end{pmatrix} \right)$$

Density of $\max\{S_{i1}\}$; under $H_0$

$$\int_{-\infty}^{\infty} \frac{k}{I_1/2} \phi\left(\frac{x}{\sqrt{I_1/2}}\right) \left\{ \Phi\left(\frac{x}{\sqrt{I_1/2}}\right) \right\}^{k-1} \phi\left(\frac{x-s}{\sqrt{I_1/2}}\right) dx$$

First look:
Replace normal density in single treatment case with density of $\max\{S_{i1}\}$

Second look:
$S_{i2} - S_{i1} \sim N(\theta(I_2 - I_1), I_2 - I_1)$ ind. of $S_{i1}$.
Hence can obtain stopping boundary to satisfy (1) and (2)

First look:
Replace normal density in single treatment case with density of $\max\{S_{i1}\}$

Second look:
$S_{i2} - S_{i1} \sim N(\theta(I_2 - I_1), I_2 - I_1)$ ind. of $S_{i1}$.
Hence can obtain stopping boundary to satisfy (1) and (2)

Critical values adjusted (upwards) to control type I error rate for selection of best-performing treatment at stage 1

First look:
Replace normal density in single treatment case with density of $\max\{S_{i1}\}$

Second look:
$S_{i2} - S_{i1} \sim N(\theta(I_2 - I_1), I_2 - I_1)$ ind. of $S_{i1}$.
Hence can obtain stopping boundary to satisfy (1) and (2)

Critical values adjusted (upwards) to control type I error rate for selection of best-performing treatment at stage 1

Type I error rate also controlled for any other selection as any $S_{i1}$ is stochastically no larger than $\max\{S_{i1}\}$

# 4. Using an early endpoint

## 4.1 Notation

Treatments: $T_0$ = control, $T_1, \ldots, T_k$ = experimental
Short- and long-term endpoints $X, Y$ for treatment $T_i$ with

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N\left(\begin{pmatrix} \nu_i \\ \mu_i \end{pmatrix}, \begin{pmatrix} \sigma_s^2 & \rho\sigma\sigma_s \\ \rho\sigma\sigma_s & \sigma^2 \end{pmatrix}\right)$$

$\theta_i = \mu_i - \mu_0$ is parameter of interest
measure of superiority of $T_i$ over $T_0$ on long-term endpoint



Stallard N. (2010) A confirmatory seamless phase II/III clinical trial design incorporating short-term endpoint information. *Stat. Med.*, **29**, 959–971.

# 4. Using an early endpoint

## 4.1 Notation

Treatments: $T_0$ = control, $T_1, \ldots, T_k$ = experimental
Short- and long-term endpoints $X, Y$ for treatment $T_i$ with

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N\left(\begin{pmatrix} \nu_i \\ \mu_i \end{pmatrix}, \begin{pmatrix} \sigma_s^2 & \rho\sigma\sigma_s \\ \rho\sigma\sigma_s & \sigma^2 \end{pmatrix}\right)$$

$\theta_i = \mu_i - \mu_0$ is parameter of interest
measure of superiority of $T_i$ over $T_0$ on long-term endpoint

Some patients have $X$ and $Y$ observed
Some patients have $X$ only observed

Stallard N. (2010) A confirmatory seamless phase II/III clinical trial design incorporating short-term endpoint information. *Stat. Med.*, **29**, 959–971.

# 4.2 Test statistic distributions

For two stages:

Stage 1

long-term data on $n_1$ per treatment

short term data on $N_1 (\geq n_1)$ per treatment

obtain $\hat{\theta}_{i1}$ and $var(\hat{\theta}_{i1}) = 2\sigma^2/n_1 - 2\sigma^2\rho^2(1/n_1 - 1/N_1)$

Select treatment with largest $\hat{\theta}_{i1}$

Stage 2

long-term data on $n_2 \geq N_1$ in total per treatment

obtain $\hat{\theta}_{i2}$ and $var(\hat{\theta}_{i2}) = 2\sigma^2/n_2$

# 4.2 Test statistic distributions

For two stages:

Stage 1

long-term data on $n_1$ per treatment

short term data on $N_1 (\geq n_1)$ per treatment

obtain $\hat{\theta}_{i1}$ and $var(\hat{\theta}_{i1}) = 2\sigma^2/n_1 - 2\sigma^2\rho^2(1/n_1 - 1/N_1)$

Select treatment with largest $\hat{\theta}_{i1}$

Stage 2

long-term data on $n_2 \geq N_1$ in total per treatment

obtain $\hat{\theta}_{i2}$ and $var(\hat{\theta}_{i2}) = 2\sigma^2/n_2$

$S_{11}, \ldots, S_{k1}, S_{i2} - S_{i1}$ have normal distributions as above

Construct critical values to control type I error rate

# 5. Example and simulation study

$k = 3$, $N_1 = 100, n_1 = 40$, $n_2 = 200$, no stopping at stage 1

| $\rho$ | Effective stage 1 $n$/gp | Standardised stage 2 critical value | Type I error rate | Power at $\theta_1 = \theta_2 = 0$ $\theta_3 = \sigma/3$) |
|---|---|---|---|---|
| 0.0 | 40 | 2.19 | 0.0256 | 0.782 |
| 0.5 | 47 | 2.20 | 0.0246 | 0.802 |
| 0.6 | 51 | 2.21 | 0.0246 | 0.801 |
| 0.7 | 57 | 2.22 | 0.0245 | 0.819 |
| 0.8 | 65 | 2.23 | 0.0249 | 0.829 |
| 0.9 | 80 | 2.25 | 0.0246 | 0.839 |

# 6. Flexible treatment selection

## 6.1 Error rate inflation ignoring flexible selection

Type I error rate is controlled provided treatment with largest $\hat{\theta}_i$ is selected

What if some other selection rule is used?

# 6. Flexible treatment selection

## 6.1 Error rate inflation ignoring flexible selection

Type I error rate is controlled provided treatment with largest $\hat{\theta}_i$ is selected

What if some other selection rule is used?

Suppose we knew (or could guess) $\nu_i$
Make selection to maximise type I error rate given all stage 1 data

Stallard N, Kunz CU, Todd S, Parsons N, Friede T. (2015) Flexible selection of a single treatment incorporating short-term endpoint information in a phase II/III clinical trial. *Stat. Med.*, **29**, 3104-3115.

Given stage 1 data

$$S_{i2} \sim N((n_2 - n_1)\theta_i + n_1\tilde{\theta}_i, 2\sigma^2(n_2 - n_1 - \rho^2(N_1 - n_1))$$

where

$$\tilde{\theta}_i = \frac{1}{n_1}\sum_{j=1}^{n_1}(Y_{ij} - Y_{0j}) + \rho\frac{\sigma}{\sigma_s}\sum_{j=n_1+1}^{N_1}(X_{ij} - X_{0j} - (\nu_i - \nu_0))$$

Type I error rate maximised by choosing treatment with largest $\tilde{\theta}_i$

| $\rho$ | Type I error | |
|---|---|---|
| | using $\hat{\theta}_i$ | using $\tilde{\theta}_i$ |
| 0.0 | 0.0256 | 0.0256 |
| 0.5 | 0.0246 | 0.0258 |
| 0.6 | 0.0246 | 0.0258 |
| 0.7 | 0.0245 | 0.0258 |
| 0.8 | 0.0249 | 0.0263 |
| 0.9 | 0.0246 | 0.0256 |

# 6.2 Correcting for flexible selection

Use joint distribution of $\max\{\tilde{\theta}_{i1}\}$ and $\hat{\theta}_{i2}$ to get critical values to control type I error rate for any selection rule

# 6.2 Correcting for flexible selection

Use joint distribution of $\max\{\tilde{\theta}_{i1}\}$ and $\hat{\theta}_{i2}$ to get critical values to control type I error rate for any selection rule

| $\rho$ | Original | | Corrected | | | |
|---|---|---|---|---|---|---|
| | Effect. $n$/gp | Power | Effect. $n$/gp | Type I error using $\tilde{\theta}_i$ | Type I error using $\hat{\theta}_i$ | Power using $\hat{\theta}_i$ |
| 0.0 | 40 | 0.782 | | | | |
| 0.5 | 47 | 0.802 | | | | |
| 0.6 | 51 | 0.810 | | | | |
| 0.7 | 57 | 0.819 | | | | |
| 0.8 | 65 | 0.829 | | | | |
| 0.9 | 80 | 0.839 | | | | |

# 6.2 Correcting for flexible selection

Use joint distribution of $\max\{\tilde{\theta}_{i1}\}$ and $\hat{\theta}_{i2}$ to get critical values to control type I error rate for any selection rule

| $\rho$ | Original | | Corrected | | | |
|---|---|---|---|---|---|---|
| | Effect. $n$/gp | Power | Effect. $n$/gp | Type I error using $\tilde{\theta}_i$ | Type I error using $\hat{\theta}_i$ | Power using $\hat{\theta}_i$ |
| 0.0 | 40 | 0.782 | 40 | 0.0256 | 0.0256 | 0.782 |
| 0.5 | 47 | 0.802 | 55 | 0.0250 | 0.0240 | 0.799 |
| 0.6 | 51 | 0.810 | 62 | 0.0246 | 0.0233 | 0.806 |
| 0.7 | 57 | 0.819 | 70 | 0.0247 | 0.0234 | 0.815 |
| 0.8 | 65 | 0.829 | 78 | 0.0250 | 0.0236 | 0.825 |
| 0.9 | 80 | 0.839 | 89 | 0.0247 | 0.0237 | 0.836 |

# 7. Conclusions and comments

New method:

enables combination of phases II and III in a single trial by
- allowing treatment selection at first interim analysis
- allowing short-term outcome data in treatment selection
- maintaining control of type I error rate

# 7. Conclusions and comments

New method:

enables combination of phases II and III in a single trial by
- allowing treatment selection at first interim analysis
- allowing short-term outcome data in treatment selection
- maintaining control of type I error rate

can incorporate early stopping

gains power through use of short-term endpoint if $\rho \neq 0$

can allow for flexible treatment selection if appropriate critical values are used