# Why I hate Minimisation

Stephen Senn

# Outline

- Why you must condition on covariates
- The true value of balance
- How to balance if you must
- Variance matters
- Why I hate minimisation

# Typical MRC Stuff

‘The central telephone randomisation system used a minimisation algorithm to balance the treatment groups with respect to eligibility criteria and other major prognostic factors.’ (p24)

‘All comparisons involved logrank analyses of the first occurrence of particular events during the scheduled treatment period after randomisation among all those allocated the vitamins versus all those allocated matching placebo capsules (ie, they were “intention-to treat” analyses).’ (p24)

1. (2002) MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20,536 high-risk individuals: a randomised placebo-controlled trial. *Lancet* **360**:7-22

# Three Games with Two Dice

- The object is to call the odds for getting a score of ten in rolling two dice
  - A red die and  a black die
- The game is played three different ways
  - Game 1. The two die are rolled together you call the odds before they are rolled
  - Game 2. The red die is rolled you are shown the score and then call the odds before the black die is rolled
  - Game 3. You call the odds. The red die is rolled first but you are not shown it and then the black one is rolled
- How should you bet?

**Table 1. Sample space for a game of chance involving two dice**

| | Red Die Score | | | | | |
|---|---|---|---|---|---|---|
| **Black Die Score** | 1 | 2 | 3 | 4 | 5 | 6 |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 4 | 5 | 6 | 7 | 8 | 9 | (*10*) |
| 5 | 6 | 7 | 8 | 9 | (*10*) | 11 |
| 6 | 7 | 8 | 9 | (*10*) | 11 | 12 |

Game 1: Probability = 3/36 = 1/12

Table 2. Probability of a total score of 10 given the red die score

| Red Die | Probability Total = 10 | Odds |
|---|---|---|
| 1 | 0 | 0 |
| 2 | 0 | 0 |
| 3 | 0 | 0 |
| 4 | 1/6 | 1:5 |
| 5 | 1/6 | 1:5 |
| 6 | 1/6 | 1:5 |

Game 2: Either the probability = 0 or the probability = 1/6

Game 3: The probability = ½ x 0 + ½ x 1/6= 1/12

# The Morals

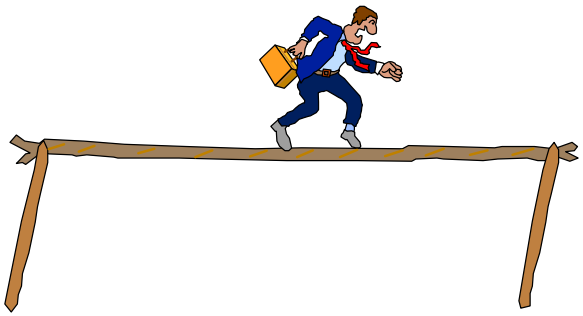
- You can't treat game 2 like game 1.
  - You must condition on the information you receive in order to be wisely
  - You must use the actual data from the red die
- You can treat game 3 like game 1.
  - You can use the *distribution in probability* that the red die has
- You can't ignore an observed prognostic covariate in analysing a clinical trial just because you randomised
  - That would be to treat game 2 like game 1
- You can ignore an unobserved covariate precisely because you did randomise
  - Because you are entitled to treat game 3 like game 1

# The Reality

Trialists continue to use their randomization as an excuse for ignoring prognostic information, and they continue to worry about the effect of factors they have not measured. Neither practice is logical.

# The True Value of Balance

- It is generally held as being self evident that a trial which is not balanced is not valid.
- Trials are examined at baseline to establish their validity.
- In fact the matter is not so simple.....



# Choices, Choices

Trial  two is balanced whereas trial one is not.

One might think that trial two provides the more reliable information.

However, the reverse is the case.

Trial one contains a comparable trial to trial two *within* it.

It is simply trial two with the addition of 8 further male patients in the verum group and 10 further female patients in the placebo group.

How could *more* information be worse than *less*?

# A Tale of two Tables

| Trial 1 | Treatment | | |
|---|---|---|---|
| Sex | Verum | Placebo | TOTAL |
| Male | 34 | 26 | 60 |
| Female | 15 | 25 | 40 |
| TOTAL | 49 | 51 | 100 |

| Trial 2 | Treatment | | |
|---|---|---|---|
| Sex | Verum | Placebo | TOTAL |
| Male | 26 | 26 | 52 |
| Female | 15 | 15 | 30 |
| TOTAL | 41 | 41 | 82 |

# Stratification

All we need to do is compare like with like.

If we compare males with males and females with females we shall obtain two unbiased estimators of the treatment effects. These can then be combined in some appropriate way. This technique is called *stratification*.

A similar approach called analysis of covariance is available to deal with continuous covariates such as height, age or a baseline measurement.

# Validity and Efficiency

- So if balance is not necessary to produce unbiased estimates what is its value?
- The answer is that it leads to efficient estimates
- If you condition on covariates the estimate will be more efficient the closer they are to being balanced
- However if you don't condition on observed balanced covariates your inferences are invalid
  - Example analysing a matched pairs design using the two-sample t-test

# What you learn in your first regression course

$$\boldsymbol{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} \quad \mathbf{X} = \begin{pmatrix} 1 & X_{11} & \cdots & X_{k1} \\ 1 & X_{12} & \ldots & X_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n} & \ldots & X_{kn} \end{pmatrix} \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} \quad \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{Y} \qquad E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta}, \quad V(\hat{\boldsymbol{\beta}}) = \sigma^2\left(\mathbf{X}'\mathbf{X}\right)^{-1}.$$

# The Value of Balance

$$\mathrm{var}(\hat{\beta}) = (X'X)^{-1}\sigma^2$$

Variance multiplier for the treatment effect

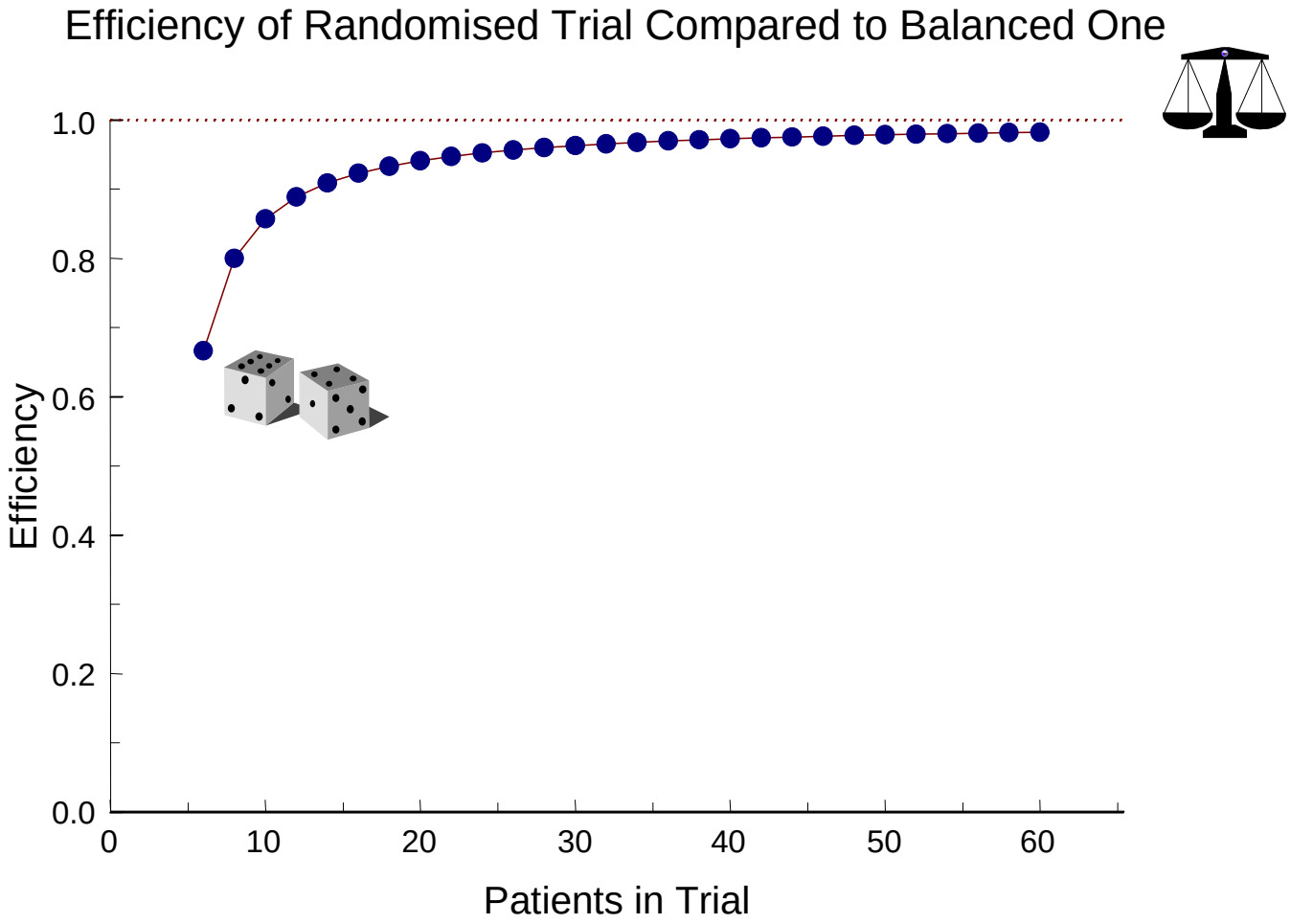
$$= \begin{pmatrix} a_{11} & a_{12} & a_{1k} \\ a_{12} & a_{22} & \\ & & \\ a_{1k} & & a_{kk} \end{pmatrix} \sigma^2$$

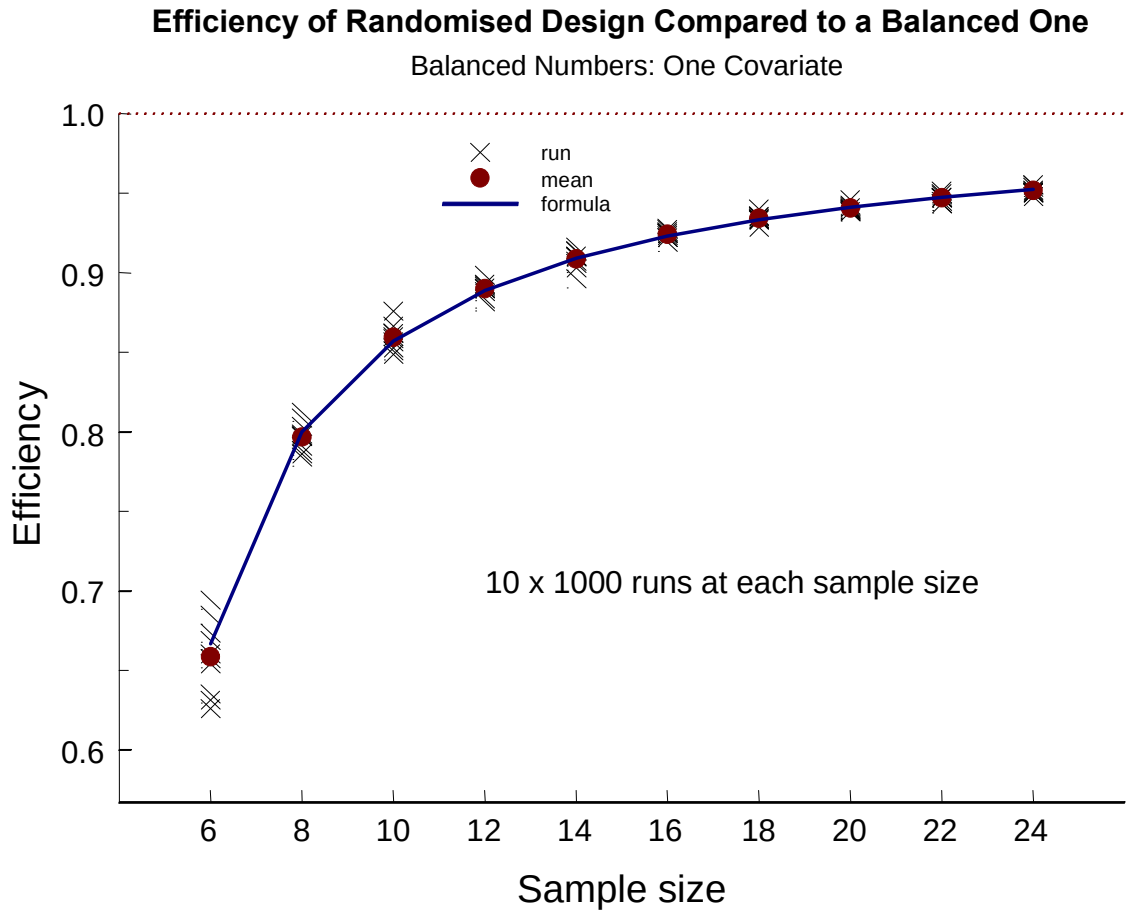
The value of $\sigma^2$ depends on the model.

The value of $a_{22}$ depends on the design and this only achieves its lower bound when covariates are balanced.

$$a_{22} \geq 2/n$$

The true value of balance


Efficiency of Randomised Trial Compared to Balanced One
Efficiency
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
50
60
Patients in Trial


Efficiency of Randomised Design Compared to a Balanced One
Balanced Numbers: One Covariate
1.0
0.9
0.8
0.7
0.6
Efficiency
run
mean
formula
10 x 1000 runs at each sample size
6
8
10
12
14
16
18
20
22
24
Sample size

# Minimisation Doesn't even Minimise Well

- In fact minimisation achieves marginal and partly irrelevant balance
  - Adding together apples and pears
- The best solution as 'enny fule  kan kno' is to work with the design matrix
- Allocate so that the variance multiplier for the treatment effect is as favourable as possible

# Atkinson's Approach

$$\left(\mathbf{X}'\mathbf{X}\right)^{-1} = \mathbf{A}$$

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k0} & a_{k1} & \cdots & a_{kk} \end{pmatrix}$$

$$\text{var}(\hat{\beta}_i) = a_{ii}\sigma^2$$

Choose allocations such that $a_{22}$ is minimised

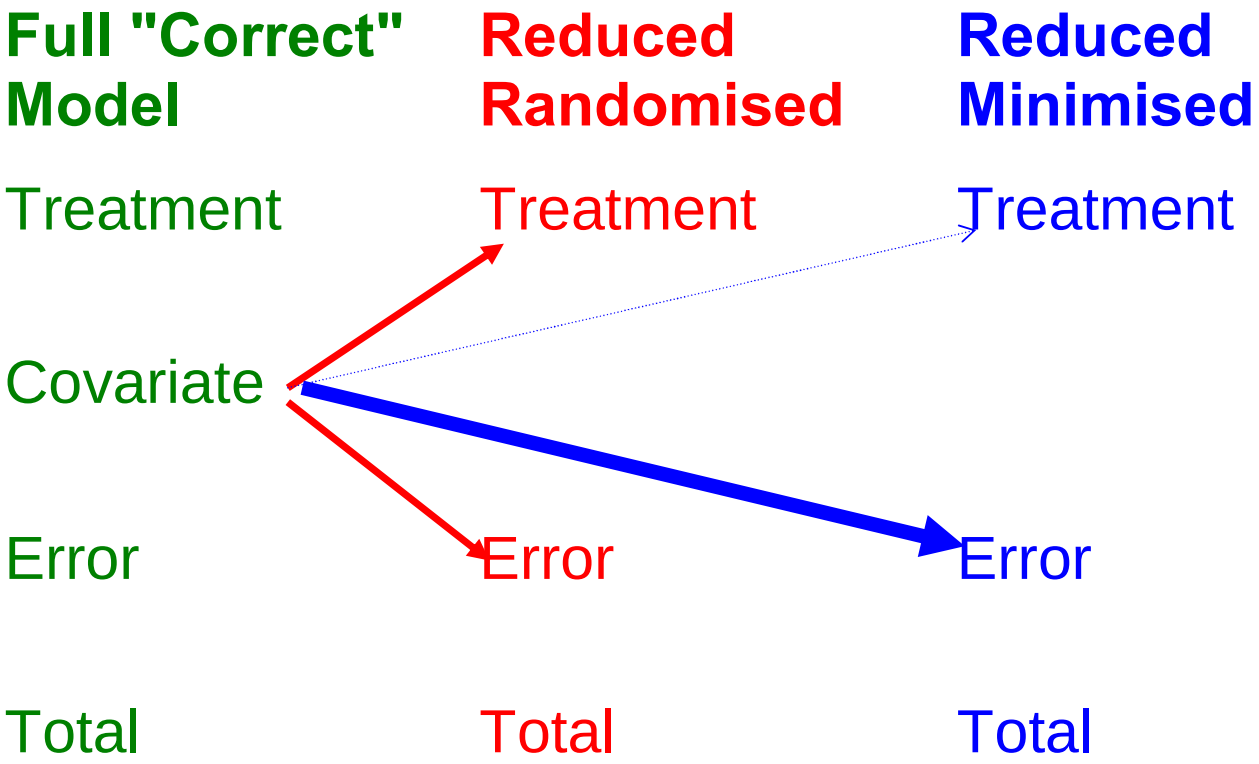
# Non-Linear Models

- This is more complicated
- Variance depends not just on design matrix but on response
- Nevertheless design matrix is important
- Furthermore the problems in not conditioning are worse
  - Gail et al, Robinson & Jewel, Ford et al

# A Problem

- All such sequential balancing methods restrict the randomisation strongly to a degree beyond that necessary to balance by the factor by the end of the trial
- This may lead to invalid variance estimates
  - incompatible with Fisher philosophy of randomised experiments
    - see also Nelder general balance
- Student's (and Taves's) argument would be that it leads to conservative inference
  - and that this is good

# Don’t Forget the Variance Estimate



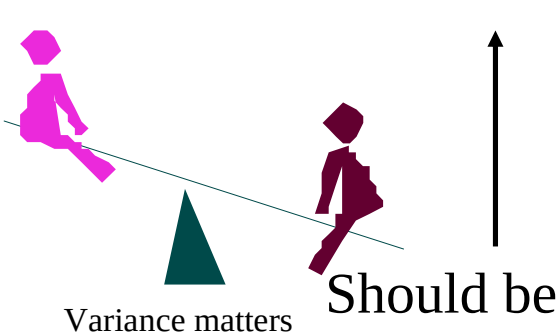
# A Red Herring

- The minimisers common defence is ‘conservative inference’
- ‘So what if our reported standard errors are higher than the true ones’
- ‘The result is *conservative inference*’
- If you like conservative inference why not just multiply all your standard errors by two?
- The next slide shows the consequence of conservative inference

# Variance, Randomisation and Meta-Analysis

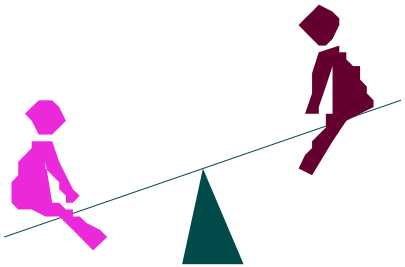
Consider the meta-analysis of two otherwise identical trials: one randomised, one balanced.

| Trial | True Variance | Weight | Est Variance | Weight |
|---|---|---|---|---|
| Randomised | Higher | Less | Lower | More |
| Balanced | Lower | More | Higher | Less |



Should be

Variance matters



Will be

# How to Eliminate The Effect of Covariates by Allocation Alone

| | A | B |
|---|---|---|
| Males | 50 | 50 |
| Females | 50 | 50 |

This eliminates the effect of sex from the unadjusted treatment difference

| | A | B |
|---|---|---|
| Males | 100 | |
| Females | | 100 |

This eliminates the effect of sex from the unadjusted within-treatment variance estimate

# How to Eliminate the Effect of a Covariate from Estimated Treatment Effect and Variance

You do this by conditioning on sex (modelling) whether or not sex and treatment are balanced

| | A | B |
|---|---|---|
| Male | 47 | 53 |
| Female | 53 | 47 |

# What About Bayesians?

- Belief dictates the model
- The model dictates the analysis
- The design determines efficiency
  - Design does ***not*** dictate analysis
- Randomised designs are (slightly) less efficient
- Why randomise?
  - Randomisation prevents your having to factor in your beliefs about how the trialists will behave

# Why I Hate Minimisation Reasons

- It is not based on sound design theory
- Its contribution to improving efficiency is minimal
- It violates randomisation based analysis
- People who use it don't even do good model-based analysis (MRC, EORTC etc balance but don't condition)
- We should fit covariates not find elaborate excuses to ignore them

# In Short

I’d rather be hanged for a sheep than a lamb

And I for one will not let the minimisers pull the wool over my eyes

# My Philosophy of Clinical Trials

- Your (reasonable) beliefs dictate the model
- You should try measure what you think is important
- You should try fit what you have measured
  - Caveat : random regressors and the Gauss-Markov theorem
- If you can balance what is important so much the better
  - But fitting is more important than balancing
- Randomisation deals with unmeasured covariates
  - You can use the distribution in probability of unmeasured covariates
  - For measured covariates you must use the actual observed distribution
- Claiming to do ‘conservative inference’ is just a convenient way of hiding bad practice

# What's out and What's in

## Out

- Log-rank test
- T-test on change scores
- Chi-square tests on 2 x 2 tables
- Responder analysis and dichotomies
- Balancing as an excuse for not conditioning

## In

- Proportional hazards
- Analysis of covariance fitting baseline
- Logistic regression fitting covariates
- Analysis of original values
- Modelling as a guide for designs

"overpaid,  oversexed and over here".

Tommy Trinder (1909-1989) on the subject of the GIs in WWII

“Over-hyped, overused and overdue for retirement”

Stephen Senn on minimisation

A plea to all right- thinking statisticians:

Help me consign this piece of garbage to the rubbish dump of history