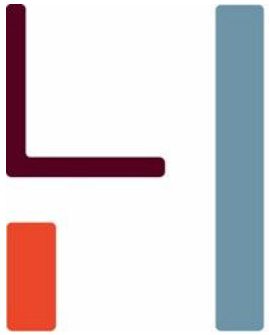
A little bit me, a little bit you

N of 1 trials, random effects and shrinkage estimators

Stephen Senn



LUXEMBOURG INSTITUTE OF **HEALTH**

RESEARCH DEDICATED TO LIFE

Acknowledgements

Many thanks to you all for the invitation and in particular to Stacy Lindborg for initiating and organizing everything

This work is partly supported by the European Union's 7th Framework Programme for research, technological development and demonstration under grant agreement no. 602552. "IDEAL"

Some of this work is joint with Artur Araujo and Sonia Leite in my group and Steven Julious at Sheffield University

Rare Diseases

- As far as the Food and Drug Administration is concerned anything that affects fewer than 200,000 people in the US
- However many diseases are much rarer than this
- But there are at least 7,000 rare diseases
- Thus the total number of persons effected is considerable

Rare Diseases Program

Mission Statement:

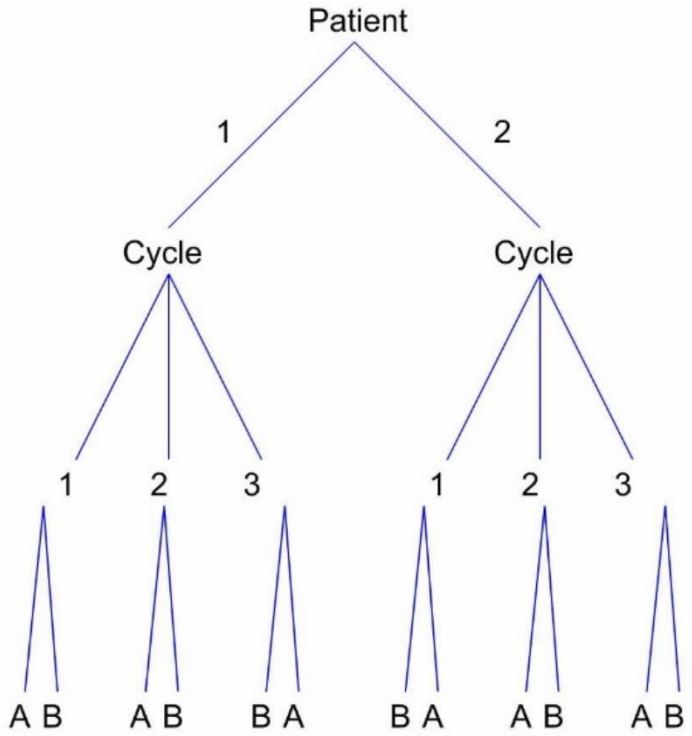
To facilitate, support and accelerate the development of drug and biologic products for the benefit of patients with rare disorders.

Overview:

- Coordinate the development of CDER policy, procedures and training for the review of treatments for rare diseases.
- Assist in outside development and maintenance of good science as the basis for the development of treatments for rare diseases.
- Work collaboratively with external and internal rare disease stakeholders to promote the development of treatments for rare disorders.
- Maintain collaborative relationships with CDER's review divisions to promote consistency and innovation in the review of treatments for rare disorders.

N-of-1 studies

- Studies in which patients are repeatedly randomised to treatment and control
- Increased efficiency because
  - Each patient acts as own control
  - More than one judgement of effect per patient
- However, only possible for chronic diseases
- Possible randomisation in *k* cycles of treatment
- Implies 2^k possible sequences



Reasons for conducting n-f-1 trials
(It is assumed that the disease is stable)

Rare disease

- Patients are few or otherwise difficult to recruit
- Within-patient studies are more efficient
- Increasing the number of periods is a way to increase the number of measurements and reduce the variance

Personalised response

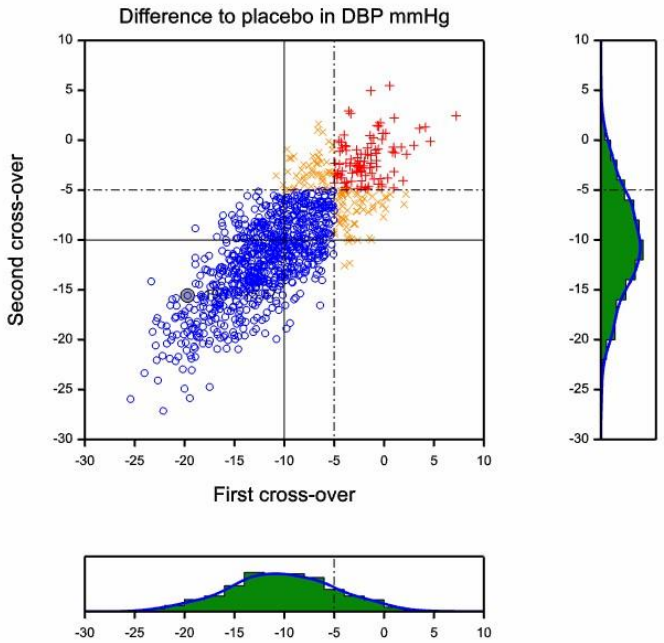
- It is desired to study personalised response to treatment
- It is necessary to separate out the components of variation
  - Within-patient
  - Treatment by patient interaction
- Designs when each patient is treated at least twice are particularly good at this

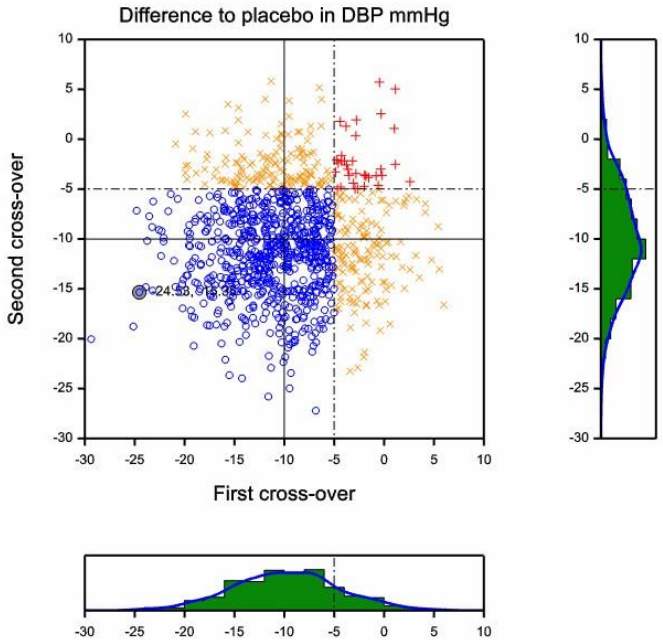
A Thought Experiment

- Imagine a cross-over trial in hypertension
- Patients randomised to receive ACE II inhibitor or placebo in random order
- Then we do it again
- Each patient does the cross-over twice
- We can compare each patient's response under ACE II to placebo twice

Design

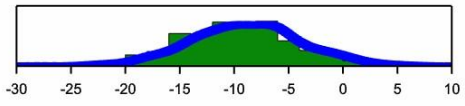
	First Cross-over		Second Cross-over	
	Period			
Sequence	1	2	3	4
I	A	B	A	B
II	B	A	B	A
III	A	B	B	A
IV	B	A	A	B





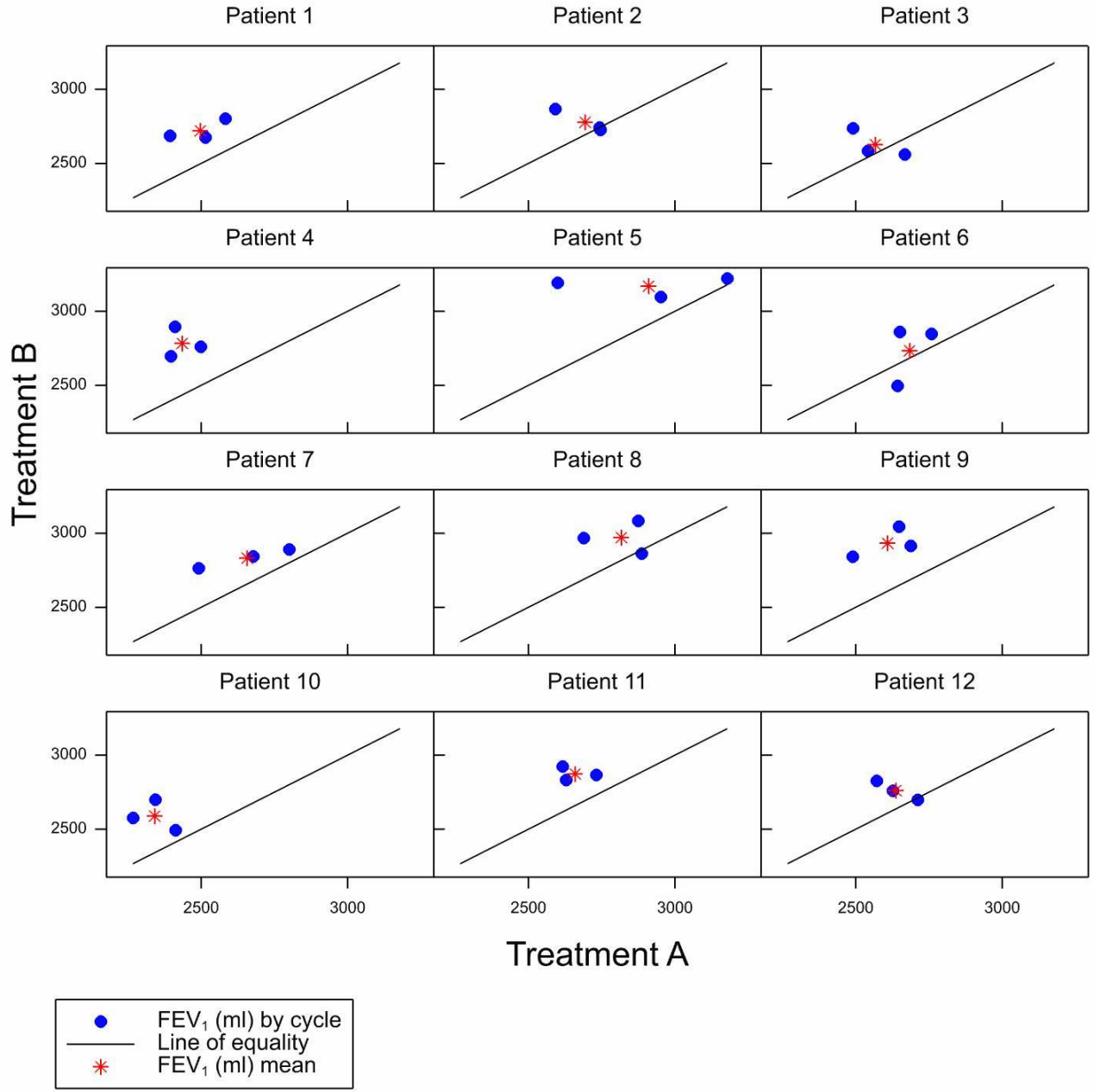
Patients are treated in two cross-over trials , thus permitting two estimates of the difference between active treatment and placebo. The difference on the second occasion is plotted against the first. Blue = response on both occasions, red = non-response on both occasions, orange = response on one occasion but not the other.

The marginal distributions are given as green histograms. LHS response on first occasion predicts response on second. RHS response on first occasion does **not** predict response on second. If you had only carried out one cross-over you would have the picture below. Which case does it apply to?



A simulated example

- Twelve patients suffering from a chronic rare respiratory complaint
  - For example cystic fibrosis
- Each patient is randomised in three pairs of periods, comparing two treatments A and B
- Adequate washout is built in to the design
- Thus we have 12 x 3 x 2 = 72 observations altogether
- Efficacy is measured using forced expiratory volume in one second (FEV_1) in ml
- How should we analyse such an experiment?


Patient 1
Patient 2
Patient 3
Patient 4
Patient 5
Patient 6
Patient 7
Patient 8
Patient 9
Patient 10
Patient 11
Patient 12
3000
2500
Treatment B
Treatment A
FEV1 (ml) by cycle
Line of equality
FEV1 (ml) mean

Possible objectives of an analysis

- **Is one of the treatments better?**
  - **Significance tests**
- What can be said about the average effect in the patients that were studied?
  - Estimates, confidence intervals
- **What can be said about the average effects in future patients?**
- **What can be said about the effect of a given patient in the trial?**
- What can be said about a future patient not in the trial?

Two different philosophies

Randomisation philosophy

- The patients in a clinical trial are taken as fixed
- The population about which inference is made is all possible randomisations
- The patients don't change, only the pattern of assignments of treatments change

Sampling philosophy

- The patients are regarded as a sample from some possible population of patients
- This is usually handled by adding error terms corresponding to various components of variance
- This approach is much more common

Is one of the treatments better? Significance tests

Rothamsted School

- Leading statisticians such as Fisher, Yates, Nelder, Bailey
- Developed analysis of variance not in terms of linear models but in terms of symmetry
- High point was John Nelder's theory of general balance (1965)

General Balance

1) Establish and define block structure
2) Establish and define treatment structure
3) Given randomisation the analysis then follows automatically

Here the block structure is

Patient/Cycle GenStat®

`Patient(Cycle)` SAS®

The treatment structure is

`Treatment`

The general balance approach

```
BLOCKSTRUCTURE Patient/Cycle
TREATMENTSTRUCTURE Treatment
ANOVA[FPROBABILITY=YES;NOMESSAGE=residual] Y
```

.

Analysis of variance

Variate: FEV_1 (mL)

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Patient stratum	11	1458791.	132617.	10.04	
Patient.Cycle stratum	24	316885.	13204.	1.04	
Patient.Cycle.*Units* stratum					
Treatment	1	641089.	641089.	50.57	<.001
Residual	35	443736.	12678.		
Total	71	2860501.			

NB This is equivalent to the matched pairs approach using the 36 cycles to provide the pairs

Comparing two models

The first is without a patient by treatment interaction

Source	DF	Type II SS	Mean Square	F Value	Pr > F
patient	11	1458791.444	132617.404	10.46	<.0001
patient*cycle	24	316884.667	13203.528	1.04	0.4479
Treatment	1	641089.389	641089.389	50.57	<.0001

Parameter	Estimate	Standard Error	t Value	Pr > |t|
mean effect	188.722222	26.5394469	7.11	<.0001

NB Analysis with proc glm of SAS®

The second is with a patient by treatment interaction

Source	DF	Type II SS	Mean Square	F Value	Pr > F
patient	11	1458791.444	132617.404	11.20	<.0001
patient*cycle	24	316884.667	13203.528	1.11	0.3960
Treatment	1	641089.389	641089.389	54.13	<.0001
patient*Treatment	11	159516.278	14501.480	1.22	0.3241

Parameter	Estimate	Standard Error	t Value	Pr > |t|
mean effect	188.722222	25.6498562	7.36	<.0001

This second approach is identical to fixed effect meta-analysis

Two more difficult questions

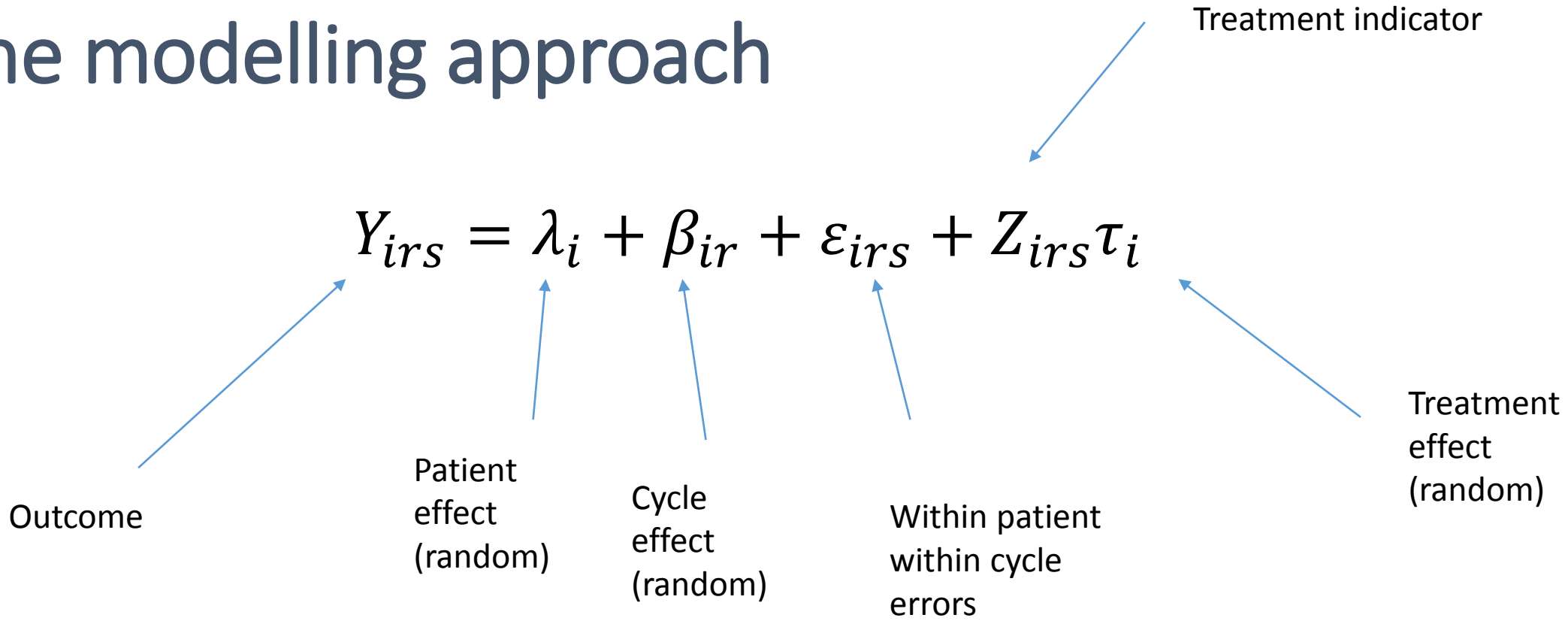
The average effects in future patients?

- This may require a mixed effects model
- Allows for a random treatment-by-patient interaction
  - The possibility that there may be variation in the effect from patient to patient
- Strong assumptions may be involved

The average effect for a given patient?

- The same random effects model can be used to predict long-term average effects for patients in the trial
- A weighted estimate is used whereby the patient's own results are averaged with the general result

The modelling approach



$$\varepsilon_{irs} \sim N(0, \sigma^2), \beta_{ir} \sim N(0, \gamma^2), \lambda_i \sim N(\Lambda, \phi^2) and\ \tau_i \sim N(T, \psi^2)$$

Shrunk estimates

Data precision

Local estimate for patient *i*

$$\hat{\tau}_i^* = \frac{\frac{1}{\psi^2}T + \frac{k}{2\sigma^2}\hat{\tau}_i}{\frac{1}{\psi^2} + \frac{k}{2\sigma^2}} = \frac{2\sigma^2 T + k\psi^2\hat{\tau}_i}{2\sigma^2 + k\psi^2}$$

Shrunk estimate

"Prior" precision

Overall treatment estimate

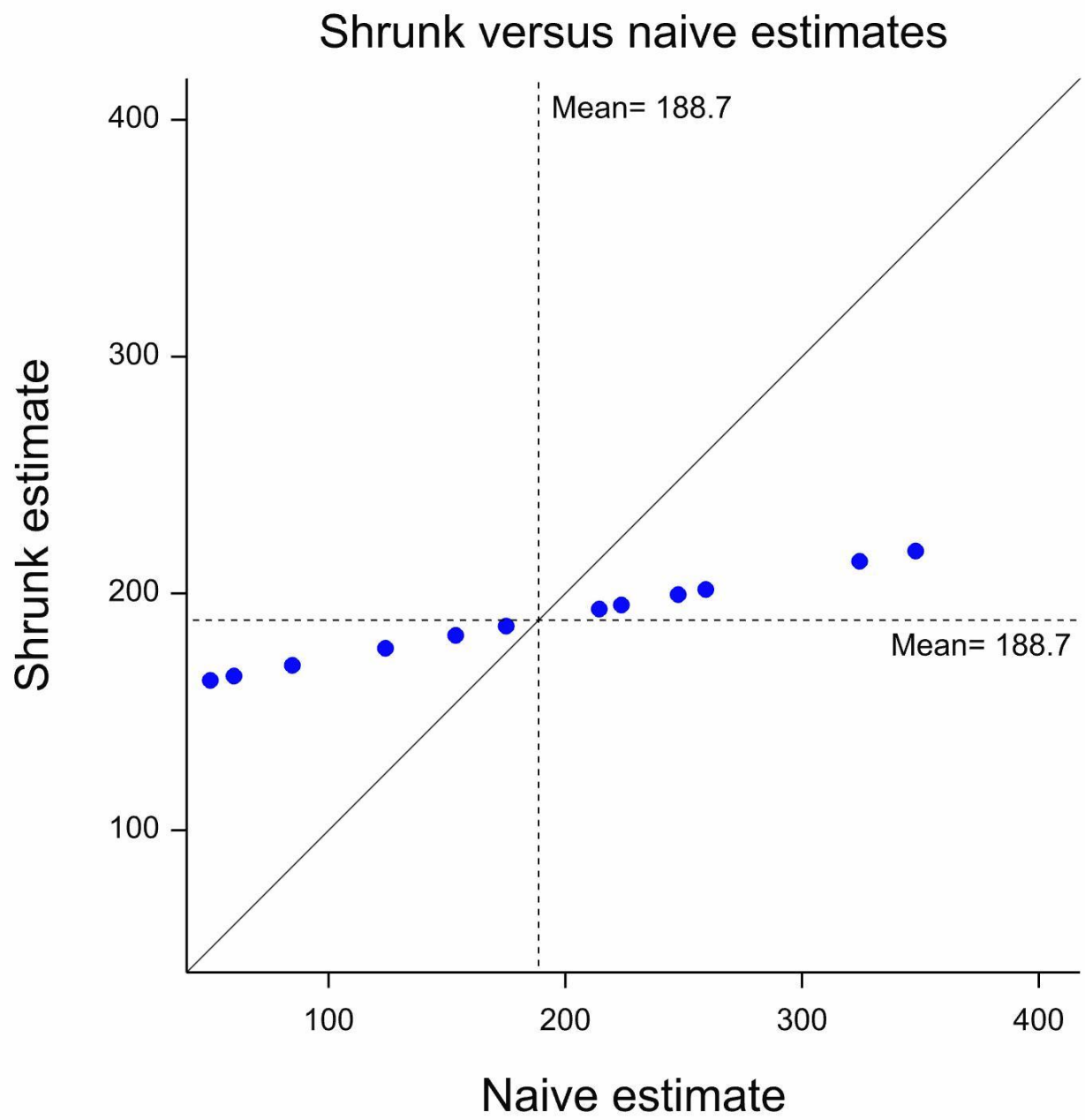
$$\frac{1}{1/\psi^2 + k/(2\sigma^2)} = \frac{2\sigma^2\psi^2}{2\sigma^2 + k\psi^2}$$

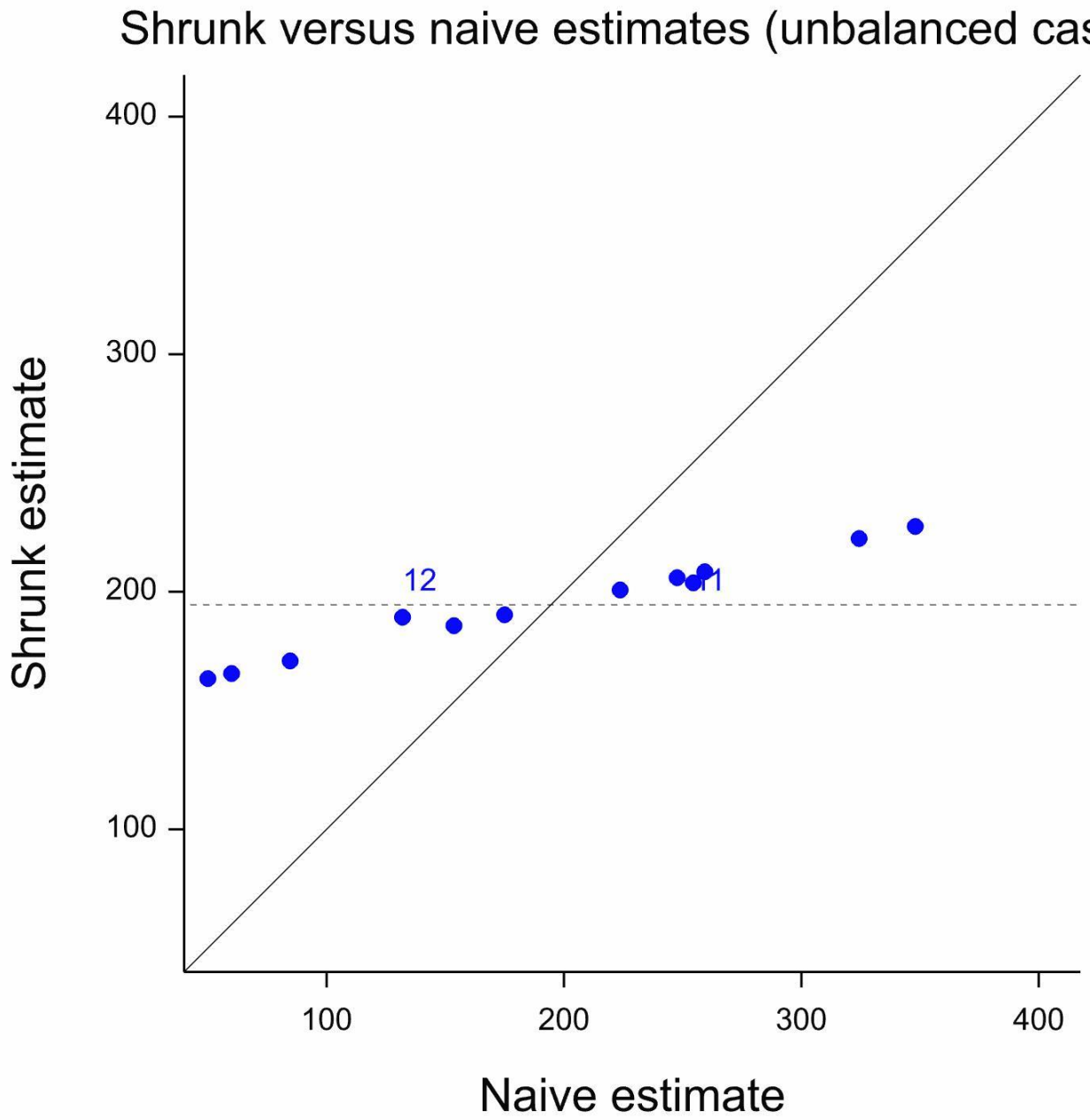
Variance

NB This is only approximately correct

Analysis using `proc mixed` of SAS®

Estimates								
Label	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
mean effect	188.72	28.3838	11	6.65	<.0001	0.05	126.25	251.19
treatment effect 1	195.13	44.5523	11	4.38	0.0011	0.05	97.0706	293.19
treatment effect 2	199.53	44.5523	11	4.48	0.0009	0.05	101.47	297.59
treatment effect 3	193.42	44.5523	11	4.34	0.0012	0.05	95.3592	291.48
treatment effect 4	176.85	44.5523	11	3.97	0.0022	0.05	78.7956	274.91
treatment effect 5	169.64	44.5523	11	3.81	0.0029	0.05	71.5834	267.70
treatment effect 6	165.12	44.5523	11	3.71	0.0035	0.05	67.0605	263.18
treatment effect 7	217.93	44.5523	11	4.89	0.0005	0.05	119.87	315.99
treatment effect 8	201.67	44.5523	11	4.53	0.0009	0.05	103.61	299.73
treatment effect 9	163.29	44.5523	11	3.67	0.0037	0.05	65.2269	261.35
treatment effect 10	186.21	44.5523	11	4.18	0.0015	0.05	88.1470	284.27
treatment effect 11	182.29	44.5523	11	4.09	0.0018	0.05	84.2353	280.35
treatment effect 12	213.59	44.5523	11	4.79	0.0006	0.05	115.53	311.65


Shrunk versus naive estimates
Mean= 188.7
Mean= 188.7
400
300
200
100
Shrunk estimate
100
200
300
400
Naive estimate



Patients 1-10 with 3 cycles
Patient 11 with 2 cycles
Patient 12 with 1 cycle

Conclusions

- Very different purposes justify very different analyses
- Proving that there is a difference between treatments (causal)
  - Randomisation based
  - Fixed effects meta-analysis
- Attempting (with difficulty) to estimate effects in patients and predict them for future patients
  - Mixed models
  - Shrinkage estimators
  - Random effects meta-analysis

Any damn fool can analyse a clinical trial and frequently does