From Rothamsted to Northwick Park: designing experiments to avoid bias and reduce variance

R. A. Bailey

Queen Mary
University of London

r.a.bailey@qmul.ac.uk

Friday Evening Discourse, Royal Institution, 27 May 2011

A few experiments

- Compare daily polypill with “do nothing” to find out which gives lower risk of stroke.

A few experiments

- Compare daily polypill with "do nothing"
  to find out which gives lower risk of stroke.
- Compare 20 varieties of wheat
  to find out which gives the most grain of bread-making quality.

A few experiments

- Compare daily polypill with "do nothing"
  to find out which gives lower risk of stroke.
- Compare 20 varieties of wheat
  to find out which gives the most grain of bread-making quality.
- Compare 3 coatings for masonry
  to find out which makes it last longest in city pollution.

A few experiments

- Compare daily polypill with “do nothing”
  to find out which gives lower risk of stroke.
- Compare 20 varieties of wheat
  to find out which gives the most grain of bread-making quality.
- Compare 3 coatings for masonry
  to find out which makes it last longest in city pollution.
- Precision measurement of the electric dipole moment of the electron, varying the levels of 9 factors (laser frequency, amplitudes of two pulses, …).

A few experiments

- Compare daily polypill with “do nothing”
  to find out which gives lower risk of stroke.
- Compare 20 varieties of wheat
  to find out which gives the most grain of bread-making quality.
- Compare 3 coatings for masonry
  to find out which makes it last longest in city pollution.
- Precision measurement of the electric dipole moment of the electron, varying the levels of 9 factors (laser frequency, amplitudes of two pulses, . . .).

A few experiments

- Compare daily polypill with "do nothing"
  to find out which gives lower risk of stroke.
- Compare 20 varieties of wheat
  to find out which gives the most grain of bread-making quality.
- Compare 3 coatings for masonry
  to find out which makes it last longest in city pollution.
- Precision measurement of the electric dipole moment of the electron, varying the levels of 9 factors (laser frequency, amplitudes of two pulses, ...).

Experiments are important in medicine, agriculture, engineering, "pure" physics, ..., and many, many areas of enquiry.

Bias: I

Suppose that we are trying to estimate an unknown number z

Bias: I

Suppose that we are trying to estimate an unknown number z (such as how much more grain of bread-making quality, in tonnes/hectare, is produced, on average, by variety A than variety B).

Bias: I

Suppose that we are trying to estimate an unknown number z (such as how much more grain of bread-making quality, in tonnes/hectare, is produced, on average, by variety A than variety B).

Our procedure is said to be unbiased if the average of all our possible estimates is the true value z.

Bias: I

Suppose that we are trying to estimate an unknown number z (such as how much more grain of bread-making quality, in tonnes/hectare, is produced, on average, by variety A than variety B).

Our procedure is said to be unbiased if the average of all our possible estimates is the true value z.

We aim to use unbiased estimation always.

Bias: I

Suppose that we are trying to estimate an unknown number z (such as how much more grain of bread-making quality, in tonnes/hectare, is produced, on average, by variety A than variety B).

Our procedure is said to be unbiased if the average of all our possible estimates is the true value z.

We aim to use unbiased estimation always.

But being right on average is not good enough …

Variance: I

We will not get exactly the same estimate if we repeat the experiment, so our estimate has some variance.

Variance: I

We will not get exactly the same estimate if we repeat the experiment, so our estimate has some variance.

The smaller the variance, the closer together are the possible estimates.

Variance: I

We will not get exactly the same estimate if we repeat the experiment, so our estimate has some variance.

The smaller the variance, the closer together are the possible estimates.

So, if our procedure is unbiased, then
the smaller the variance, the closer is our estimate to the true value z (usually).

Variance: I

We will not get exactly the same estimate if we repeat the experiment, so our estimate has some variance.

The smaller the variance, the closer together are the possible estimates.

So, if our procedure is unbiased, then
the smaller the variance, the closer is our estimate to the true value z (usually).

In fact, if our procedure is unbiased, the variance is V and our estimated value is e, then

$$e - 3\sqrt{V} \leq z \leq e + 3\sqrt{V}$$

almost all the time.

Variance: I

We will not get exactly the same estimate if we repeat the experiment, so our estimate has some variance.

The smaller the variance, the closer together are the possible estimates.

So, if our procedure is unbiased, then
the smaller the variance, the closer is our estimate to the true value z (usually).

In fact, if our procedure is unbiased, the variance is V and our estimated value is e, then

$$e - 3\sqrt{V} \leq z \leq e + 3\sqrt{V}$$

almost all the time.

The smaller the variance, the closer is our estimate to the true value.

Variance: I

We will not get exactly the same estimate if we repeat the experiment, so our estimate has some variance.

The smaller the variance, the closer together are the possible estimates.

So, if our procedure is unbiased, then
the smaller the variance, the closer is our estimate to the true value z (usually).

In fact, if our procedure is unbiased, the variance is V and our estimated value is e, then

$$e - 3\sqrt{V} \leq z \leq e + 3\sqrt{V}$$

almost all the time.

The smaller the variance, the closer is our estimate to the true value.

We aim to make variance small.

Some criteria for designing an experiment

- remove bias

Some criteria for designing an experiment

- remove bias
- make variance as small as possible

Some criteria for designing an experiment

- remove bias
- make variance as small as possible
- stay within constraints of cost, feasibility, ...,

Some criteria for designing an experiment

- remove bias
- make variance as small as possible
- stay within constraints of cost, feasibility, . . . ,

Some criteria for designing an experiment

- remove bias
- make variance as small as possible
- stay within constraints of cost, feasibility, . . . , but an experiment too small to find out anything may be a waste of resources.

Some criteria for designing an experiment

- remove bias
- make variance as small as possible
- stay within constraints of cost, feasibility, . . . , but an experiment too small to find out anything may be a waste of resources.

Why does this matter?

Some criteria for designing an experiment

- remove bias
- make variance as small as possible
- stay within constraints of cost, feasibility, . . . , but an experiment too small to find out anything may be a waste of resources.

Why does this matter?

Better quality experiments enable us to make better quality decisions to make better use of Earth's resources and to save lives.

Bias II: randomization

One way to avoid bias is to randomize:

Bias II: randomization

One way to avoid bias is to randomize:
write down a systematic plan then permute it by a randomly-chosen permutation.

Lanarkshire milk experiment: early 20th century

Treatments: extra milk rations or not.
These should have been randomized to the children within each school.
The teachers decided to give the extra milk rations to those children who were most undernourished.

Rothamsted Experimental Station (Harpenden)

This was founded by Sir John Bennet Lawes in 1843.

Broadbalk

Rothamsted Experimental Station (Harpenden)

This was founded by Sir John Bennet Lawes in 1843.

trees →

Broadbalk

Rothamsted Experimental Station (Harpenden)

This was founded by Sir John Bennet Lawes in 1843.

trees →

Broadbalk

I worked in the Statistics Department there from 1981 to 1990.

An experiment at Rothamsted that I designed

Variance II: replication

Suppose that we have N plots available and want to compare varieties A and B.

Variance II: replication

Suppose that we have N plots available and want to compare varieties A and B.

If variety A is planted on n plots, and variety B is planted on m plots, where $n+m=N$, and the variance of each yield is σ^2, then the variance of the estimate of the difference between A and B is

$$\sigma^2\left(\frac{1}{n}+\frac{1}{m}\right)=\sigma^2\left(\frac{n+m}{nm}\right)=\sigma^2\left(\frac{N}{nm}\right).$$

Variance II: replication

Suppose that we have N plots available and want to compare varieties A and B.

If variety A is planted on n plots, and variety B is planted on m plots, where $n+m=N$, and the variance of each yield is σ^2, then the variance of the estimate of the difference between A and B is

$$\sigma^2\left(\frac{1}{n}+\frac{1}{m}\right)=\sigma^2\left(\frac{n+m}{nm}\right)=\sigma^2\left(\frac{N}{nm}\right).$$

Theorem

If the total $n+m$ is fixed, the value of $\dfrac{1}{nm}$ is smallest when $|m-n|\leq 1$.

Variance III: a proof

Theorem

If the total $n+m$ is fixed, the value of $\frac{1}{nm}$ is smallest when $|m-n| \leq 1$.

Variance III: a proof

Theorem

If the total $n+m$ is fixed, the value of $\frac{1}{nm}$ is smallest when $|m-n| \leq 1$.

Proof.

Consider changing m to $m-1$ and n to $n+1$.

Variance III: a proof

Theorem

If the total $n+m$ is fixed, the value of $\frac{1}{nm}$ is smallest when $|m-n| \leq 1$.

Proof.

Consider changing m to $m-1$ and n to $n+1$.

$$\begin{aligned}
\text{new variance is smaller} \quad &\iff \quad \frac{1}{(n+1)(m-1)} < \frac{1}{nm} \\
&\iff \quad (n+1)(m-1) > nm \\
&\iff \quad nm + m - n - 1 > nm \\
&\iff \quad m - n > 1.
\end{aligned}$$

Variance III: a proof

Theorem

If the total $n+m$ is fixed, the value of $\frac{1}{nm}$ is smallest when $|m-n| \leq 1$.

Proof.

Consider changing m to $m-1$ and n to $n+1$.

$$\begin{aligned}
\text{new variance is smaller} &\iff \frac{1}{(n+1)(m-1)} < \frac{1}{nm} \\
&\iff (n+1)(m-1) > nm \\
&\iff nm + m - n - 1 > nm \\
&\iff m - n > 1.
\end{aligned}$$

If $m-n \geq 2$ (or $n-m \geq 2$), we can change the replications to get a design with smaller variance. □

Variance IV: many varieties

If we have varieties $1, 2, \ldots, v$,
then we want to minimize the average of the variance of the estimate of the difference between varieties i and j, for $1 \leq i < j \leq v$.

Variance IV: many varieties

If we have varieties 1, 2, …, v,
then we want to minimize the average of the variance of the estimate of the difference between varieties i and j, for $1 \leq i < j \leq v$.

This is achieved by making all the replications as equal as possible.

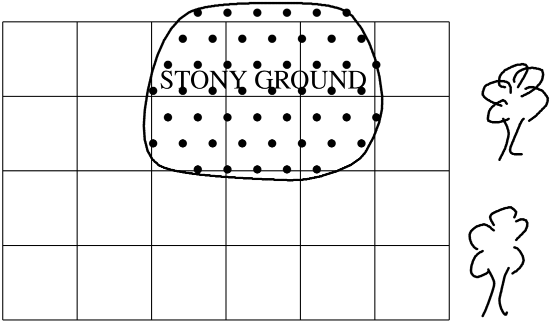
Variance IV: many varieties

If we have varieties 1, 2, $\ldots$, v,
then we want to minimize the average of the variance of the estimate of the difference between varieties i and j, for $1 \leq i < j \leq v$.

This is achieved by making all the replications as equal as possible.

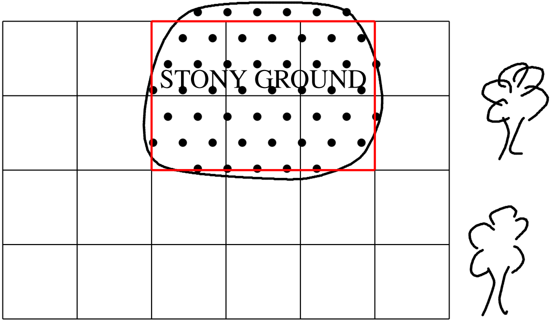
Blocking

We have 6 varieties to compare in this field. How do we avoid bias?



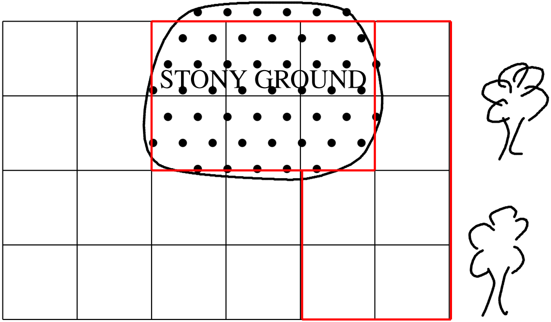
Blocking

We have 6 varieties to compare in this field. How do we avoid bias?



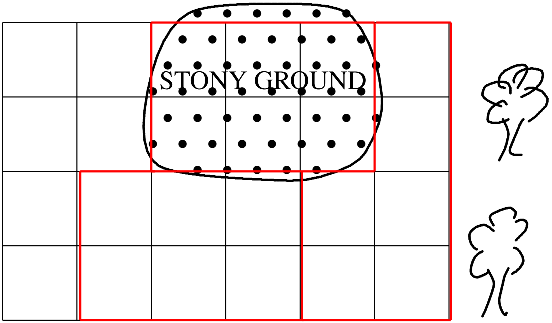
Blocking

We have 6 varieties to compare in this field. How do we avoid bias?



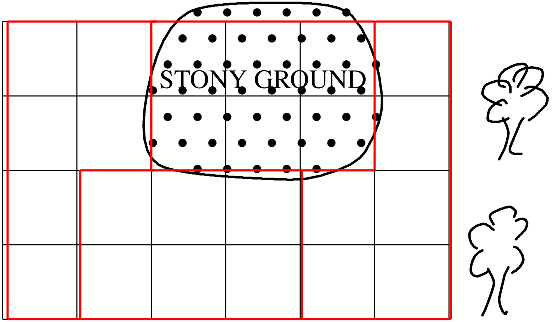
Blocking

We have 6 varieties to compare in this field. How do we avoid bias?



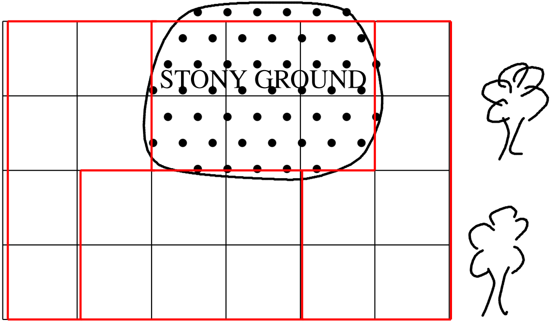
Blocking

We have 6 varieties to compare in this field. How do we avoid bias?



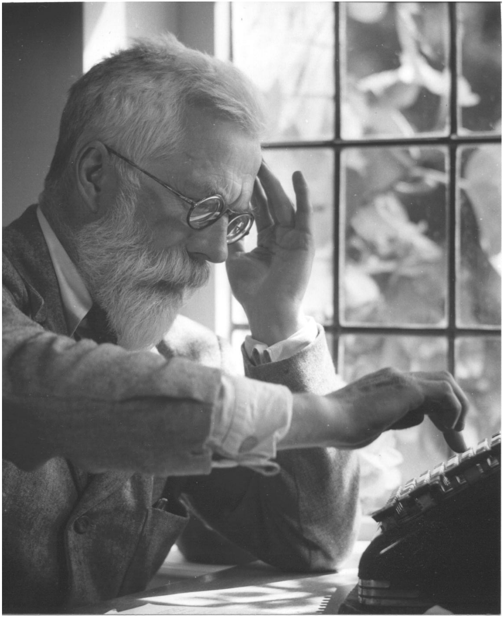
Blocking

We have 6 varieties to compare in this field. How do we avoid bias?



Partition the experimental units into homogeneous blocks
and apply each treatment to one plot in each block.

R. A. Fisher, statistician at Rothamsted 1919–1933



- randomization
- replication
- blocking

1952 portrait by Barrington Brown, reproduced by permission of the Fisher Memorial Trust

Incomplete blocks

What do we do if the blocks are too small for each one to contain all the treatments?

Incomplete blocks

What do we do if the blocks are too small for each one to contain all the treatments?

Then the unbiased estimates with the smallest variance are no longer the differences between the simple treatment means. There is a complicated formula for the average pairwise variance. It depends on the design as well as on the replications.

Incomplete blocks

What do we do if the blocks are too small for each one to contain all the treatments?

Then the unbiased estimates with the smallest variance are no longer the differences between the simple treatment means. There is a complicated formula for the average pairwise variance. It depends on the design as well as on the replications.

A design for v treatments in b blocks of size k is balanced if there is some constant λ such that every pair of treatments occur together in precisely λ blocks.

Two designs with $v = 7$, $b = 7$, $k = 3$: columns are blocks

1	2	3	4	5	6	7
2	3	4	5	6	7	1
4	5	6	7	1	2	3

balanced ($\lambda = 1$)

1	2	3	4	5	6	7
2	3	4	5	6	7	1
3	4	5	6	7	1	2

non-balanced

Results about balanced incomplete-block designs

v = number of treatments b = number of blocks
k = block size

Theorem

1. *In a BIBD,*

Results about balanced incomplete-block designs

$v =$ number of treatments $\qquad b =$ number of blocks
$k =$ block size

Theorem

1. *In a BIBD,*
   1.1 *every treatment occurs in r blocks, where $vr = bk$;*

Results about balanced incomplete-block designs

$v =$ number of treatments $\qquad b =$ number of blocks
$k =$ block size

Theorem

1. *In a BIBD,*
   1.1 *every treatment occurs in r blocks, where $vr = bk$;*
   1.2 $r(k-1) = (v-1)\lambda$*;*

Results about balanced incomplete-block designs

v = number of treatments b = number of blocks
k = block size

Theorem

1. *In a BIBD,*
   1.1 *every treatment occurs in r blocks, where $vr = bk$;*
   1.2 $r(k-1) = (v-1)\lambda$;
   1.3 $v \leq b$.

Results about balanced incomplete-block designs

v = number of treatments $\qquad$ b = number of blocks
k = block size

Theorem

1. *In a BIBD,*
   1.1 *every treatment occurs in r blocks, where $vr = bk$;*
   1.2 $r(k-1) = (v-1)\lambda$;
   1.3 $v \leq b$.
2. *BIBDs do not exist for all values of v, b and k.*

Results about balanced incomplete-block designs

$v =$ number of treatments $\qquad b =$ number of blocks
$k =$ block size

Theorem

1. *In a BIBD,*
   1.1 *every treatment occurs in r blocks, where $vr = bk$;*
   1.2 $r(k-1) = (v-1)\lambda$*;*
   1.3 $v \leq b$.
2. *BIBDs do not exist for all values of v, b and k.*
3. *If there is a BIBD, then it gives the minimum average variance of pairwise differences.*

Kirkman's Schoolgirls Problem (1847)

There are 15 schoolgirls in a certain class.
Every day, they go for a walk, and the teacher insists that they walk in groups of size 3.
Arrange the girls in groups for a week (7 days) in such a way that each pair of girls walk together in a group exactly once.

Kirkman's Schoolgirls Problem (1847)

There are 15 schoolgirls in a certain class.
Every day, they go for a walk, and the teacher insists that they walk in groups of size 3.
Arrange the girls in groups for a week (7 days) in such a way that each pair of girls walk together in a group exactly once.

This is a BIBD with $v = 15$, $b = 5 \times 7 = 35$ and $k = 3$,
with the extra property that there are five whole groups per day.

Kirkman's Schoolgirls Problem (1847)

There are 15 schoolgirls in a certain class.
Every day, they go for a walk, and the teacher insists that they walk in groups of size 3.
Arrange the girls in groups for a week (7 days) in such a way that each pair of girls walk together in a group exactly once.

This is a BIBD with $v = 15$, $b = 5 \times 7 = 35$ and $k = 3$,
with the extra property that there are five whole groups per day.

BIBDs have been studied extensively by pure mathematicians as well as statisticians.

Kirkman's Schoolgirls Problem (1847)

There are 15 schoolgirls in a certain class.
Every day, they go for a walk, and the teacher insists that they walk in groups of size 3.
Arrange the girls in groups for a week (7 days) in such a way that each pair of girls walk together in a group exactly once.

This is a BIBD with $v = 15$, $b = 5 \times 7 = 35$ and $k = 3$,
with the extra property that there are five whole groups per day.

BIBDs have been studied extensively by pure mathematicians as well as statisticians.

Homework

Solve Kirkman's Problem for 15 *schoolgirls.*

From Rothamsted to London

In 1991 I left Rothamsted and joined the University of London.

I have continued to help with the design of experiments in many areas, such as

- human–computer interaction
- biomaterials
- two-phase variety trials
- biodiversity in freshwater systems
- genomics
- a cross-over grazing trial
- the effect of plant spacing on insect populations.

New Delhi, December 2006

Experiments in microarrays

At the end of the 1990s, there was an explosion of genomic experiments using microarrays.

With some over-simplification,
these are block designs with all blocks of size 2.

Experiments in microarrays

At the end of the 1990s, there was an explosion of genomic experiments using microarrays.

With some over-simplification,
these are block designs with all blocks of size 2.

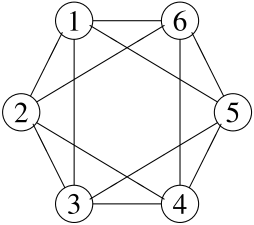
So which designs have the smallest average variance of the estimates of pairwise differences?

Some designs for 6 treatments in blocks of size 2

12 blocks (edges)

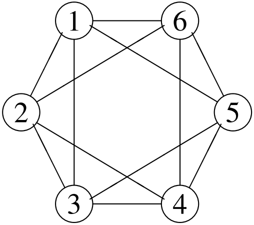
Some designs for 6 treatments in blocks of size 2

12 blocks (edges)



Some designs for 6 treatments in blocks of size 2

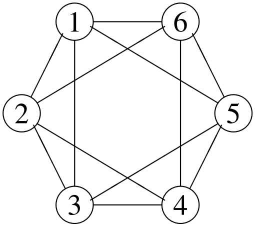
12 blocks (edges)



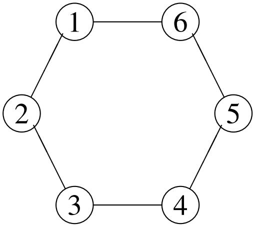
6 blocks (edges)

Some designs for 6 treatments in blocks of size 2

12 blocks (edges)

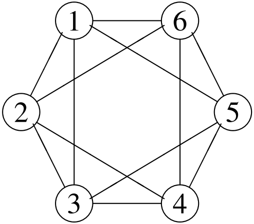


6 blocks (edges)

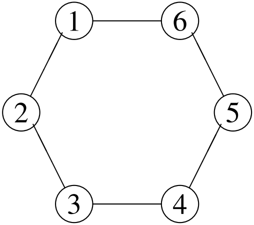


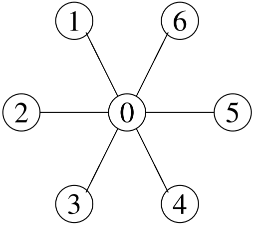
Some designs for 6 treatments in blocks of size 2

12 blocks (edges)



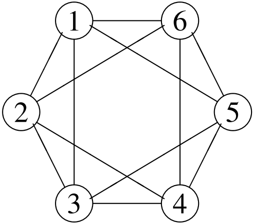
6 blocks (edges)

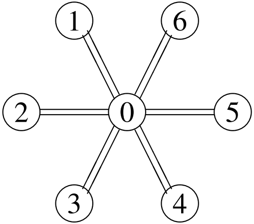




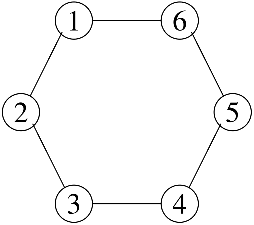
Some designs for 6 treatments in blocks of size 2

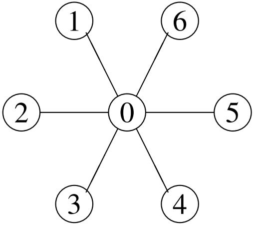
12 blocks (edges)





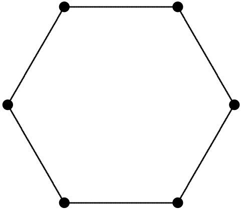
6 blocks (edges)



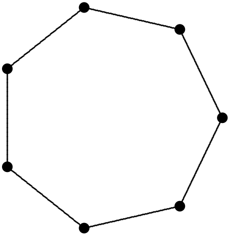


Designs with smallest variance when $k = 2$ and $b = v$

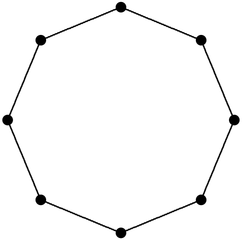
$v = 6$



$v = 7$

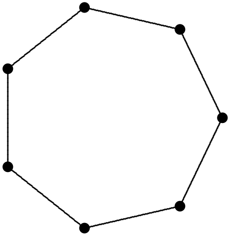


$v = 8$

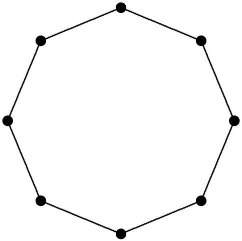


Designs with smallest variance when $k = 2$ and $b = v$

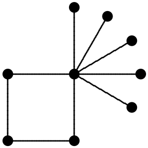
$v = 7$



$v = 8$

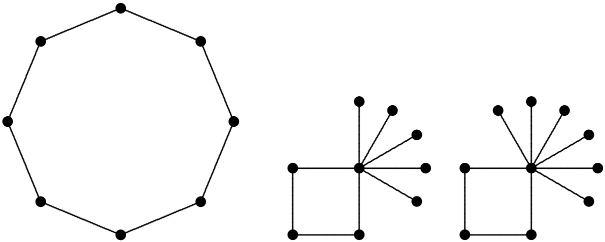


$v = 9$



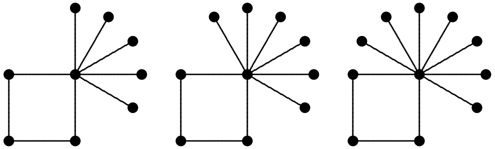
Designs with smallest variance when $k = 2$ and $b = v$

$v = 8$ $v = 9$ $v = 10$



Designs with smallest variance when $k = 2$ and $b = v$

$v = 9$ $\quad$ $v = 10$ $\quad$ $v = 11$



Designs with average replication less than 2.5

Which designs have the smallest variance
when average replication is low,
for arbitrary block size?

Research is still ongoing.

There are many strange results.

Northwick Park: the TeGenero trial

First-in-Man trial of a monoclonal antibody on healthy volunteers, March 2006: 4 cohorts of 8 volunteers each.

Cohort	TGN1412		Placebo
	Dose mg/kg body-weight	Number of Subjects	Number of Subjects
1	0.1	6	2
2	0.5	6	2
3	2.0	6	2
4	5.0	6	2

What happened to Cohort 1 on 13 March 2006

Healthy Volunteer	Randomized to	Time of intravenous administration	Time of transfer to critical care
A	TGN1412 8.4mg	0800	2400
B	Placebo	0810	
C	TGN1412 6.8mg	0820	2350
D	TGN1412 8.8mg	0830	0030
E	TGN1412 8.2mg	0840	2040
F	TGN1412 7.2mg	0850	0050
G	TGN1412 8.2mg	0900	0100
H	Placebo	0910	

The Royal Statistical Society's Working Party on Statistical Issues in First-in-Man Studies: Membership

Dipti Amin, Senior Vice-President, Quintiles
R. A. Bailey, Professor of Statistics, QMUL
Sheila Bird, Principal Scientist/Statistician, MRC Biostatistics Unit
Barbara Bogacka, Reader in Probability and Statistics, QMUL
Peter Colman, Senior Consultant Statistician, Pfizer
Andrew Garrett, Vice-President Statistics, Quintiles
Andrew Grieve, Professor of Medical Statistics, KCL
Peter Lachmann, FRS, Emeritus Professor of Immunology, Cambridge
Stephen Senn, Professor of Statistics, Glasgow

The Royal Statistical Society's Working Party on Statistical Issues in First-in-Man Studies: Membership

Dipti Amin, Senior Vice-President, Quintiles
R. A. Bailey, Professor of Statistics, QMUL
Sheila Bird, Principal Scientist/Statistician, MRC Biostatistics Unit
Barbara Bogacka, Reader in Probability and Statistics, QMUL
Peter Colman, Senior Consultant Statistician, Pfizer
Andrew Garrett, Vice-President Statistics, Quintiles
Andrew Grieve, Professor of Medical Statistics, KCL
Peter Lachmann, FRS, Emeritus Professor of Immunology, Cambridge
Stephen Senn, Professor of Statistics, Glasgow

What does block mean?

The Royal Statistical Society's Working Party on Statistical Issues in First-in-Man Studies: Membership

Dipti Amin, Senior Vice-President, Quintiles
R. A. Bailey, Professor of Statistics, QMUL
Sheila Bird, Principal Scientist/Statistician, MRC Biostatistics Unit
Barbara Bogacka, Reader in Probability and Statistics, QMUL
Peter Colman, Senior Consultant Statistician, Pfizer
Andrew Garrett, Vice-President Statistics, Quintiles
Andrew Grieve, Professor of Medical Statistics, KCL
Peter Lachmann, FRS, Emeritus Professor of Immunology, Cambridge
Stephen Senn, Professor of Statistics, Glasgow

What does block mean? strata?

The Royal Statistical Society's Working Party on Statistical Issues in First-in-Man Studies: Membership

Dipti Amin, Senior Vice-President, Quintiles
R. A. Bailey, Professor of Statistics, QMUL
Sheila Bird, Principal Scientist/Statistician, MRC Biostatistics Unit
Barbara Bogacka, Reader in Probability and Statistics, QMUL
Peter Colman, Senior Consultant Statistician, Pfizer
Andrew Garrett, Vice-President Statistics, Quintiles
Andrew Grieve, Professor of Medical Statistics, KCL
Peter Lachmann, FRS, Emeritus Professor of Immunology, Cambridge
Stephen Senn, Professor of Statistics, Glasgow

What does block mean? strata? randomize?

The Royal Statistical Society's Working Party on Statistical Issues in First-in-Man Studies: Report

Published free-standing and online in March 2007, then in *Journal of the Royal Statistical Society, Series A* **170** (2007), 517–579.

Recommendations include

- generic issues

The Royal Statistical Society's Working Party on Statistical Issues in First-in-Man Studies: Report

Published free-standing and online in March 2007, then in *Journal of the Royal Statistical Society, Series A* **170** (2007), 517–579.

Recommendations include

- generic issues
- risk (quantification; novel type of medicine; public debate)

The Royal Statistical Society's Working Party on Statistical Issues in First-in-Man Studies: Report

Published free-standing and online in March 2007, then in *Journal of the Royal Statistical Society, Series A* **170** (2007), 517–579.

Recommendations include

- generic issues
- risk (quantification; novel type of medicine; public debate)
- sharing information on adverse events (usable database)

The Royal Statistical Society's Working Party on Statistical Issues in First-in-Man Studies: Report

Published free-standing and online in March 2007, then in *Journal of the Royal Statistical Society, Series A* **170** (2007), 517–579.

Recommendations include

- generic issues
- risk (quantification; novel type of medicine; public debate)
- sharing information on adverse events (usable database)
- proper interval between dosing subjects
  (sudden adverse effects → do not dose further subjects;
  delayed adverse effects → ill subjects can be treated one by one)

The Royal Statistical Society's Working Party on Statistical Issues in First-in-Man Studies: Report

Published free-standing and online in March 2007, then in *Journal of the Royal Statistical Society, Series A* **170** (2007), 517–579.

Recommendations include

- generic issues
- risk (quantification; novel type of medicine; public debate)
- sharing information on adverse events (usable database)
- proper interval between dosing subjects
  (sudden adverse effects → do not dose further subjects;
  delayed adverse effects → ill subjects can be treated one by one)
- preclinical / clinical interface

The Royal Statistical Society's Working Party on Statistical Issues in First-in-Man Studies: Report

Published free-standing and online in March 2007, then in *Journal of the Royal Statistical Society, Series A* **170** (2007), 517–579.

Recommendations include

- generic issues
- risk (quantification; novel type of medicine; public debate)
- sharing information on adverse events (usable database)
- proper interval between dosing subjects
  (sudden adverse effects → do not dose further subjects;
  delayed adverse effects → ill subjects can be treated one by one)
- preclinical / clinical interface
- protocol

The Royal Statistical Society's Working Party on Statistical Issues in First-in-Man Studies: Report

Published free-standing and online in March 2007, then in *Journal of the Royal Statistical Society, Series A* **170** (2007), 517–579.

Recommendations include

- generic issues
- risk (quantification; novel type of medicine; public debate)
- sharing information on adverse events (usable database)
- proper interval between dosing subjects
  (sudden adverse effects → do not dose further subjects;
  delayed adverse effects → ill subjects can be treated one by one)
- preclinical / clinical interface
- protocol
- sequential choice of dose

The Royal Statistical Society's Working Party on Statistical Issues in First-in-Man Studies: Report

Published free-standing and online in March 2007, then in *Journal of the Royal Statistical Society, Series A* **170** (2007), 517–579.

Recommendations include

- generic issues
- risk (quantification; novel type of medicine; public debate)
- sharing information on adverse events (usable database)
- proper interval between dosing subjects
  (sudden adverse effects → do not dose further subjects;
  delayed adverse effects → ill subjects can be treated one by one)
- preclinical / clinical interface
- protocol
- sequential choice of dose
- allocation of ordinal doses to cohorts.

The Royal Statistical Society's Working Party on Statistical Issues in First-in-Man Studies: Report

Published free-standing and online in March 2007, then in *Journal of the Royal Statistical Society, Series A* **170** (2007), 517–579.

Recommendations include

- generic issues
- risk (quantification; novel type of medicine; public debate)
- sharing information on adverse events (usable database)
- proper interval between dosing subjects
  (sudden adverse effects → do not dose further subjects;
  delayed adverse effects → ill subjects can be treated one by one)
- preclinical / clinical interface
- protocol
- sequential choice of dose
- allocation of ordinal doses to cohorts.

Planned analysis of the TeGenero trial

Cohort	TGN1412		Placebo
	Dose	Number	Number
1	1	6	2
2	2	6	2
3	3	6	2
4	4	6	2

If all responses are uncorrelated with variance σ^2 then
Variance (dose $i -$ placebo) in cohort i is $\left(\frac{1}{6}+\frac{1}{2}\right)\sigma^2 = \frac{2}{3}\sigma^2$

Planned analysis of the TeGenero trial

Cohort	TGN1412		Placebo
	Dose	Number	Number
1	1	6	2
2	2	6	2
3	3	6	2
4	4	6	2

If all responses are uncorrelated with variance σ^2 then
Variance (dose $i -$ placebo) in cohort i is $\left(\frac{1}{6}+\frac{1}{2}\right)\sigma^2 = \frac{2}{3}\sigma^2$

From the protocol: "data of subjects having received placebo will be pooled in one group for analyses."

Planned analysis of the TeGenero trial

Cohort	TGN1412		Placebo
	Dose	Number	Number
1	1	6	2
2	2	6	2
3	3	6	2
4	4	6	2

If all responses are uncorrelated with variance σ^2 then
Variance (dose $i -$ placebo) in cohort i is $\left(\frac{1}{6}+\frac{1}{2}\right)\sigma^2 = \frac{2}{3}\sigma^2$

From the protocol: "data of subjects having received placebo will be pooled in one group for analyses."

Variance (dose $i -$ placebo) is $\left(\frac{1}{6}+\frac{1}{8}\right)\sigma^2 = \frac{7}{24}\sigma^2$

Planned analysis of the TeGenero trial

Cohort	TGN1412		Placebo
	Dose	Number	Number
1	1	6	2
2	2	6	2
3	3	6	2
4	4	6	2

If all responses are uncorrelated with variance σ^2 then
Variance (dose $i -$ placebo) in cohort i is $\left(\frac{1}{6}+\frac{1}{2}\right)\sigma^2 = \frac{2}{3}\sigma^2$

From the protocol: "data of subjects having received placebo will be pooled in one group for analyses."

Variance (dose $i -$ placebo) is $\left(\frac{1}{6}+\frac{1}{8}\right)\sigma^2 = \frac{7}{24}\sigma^2$ if there are no cohort effects.

Planned analysis of the TeGenero trial

Cohort	TGN1412		Placebo
	Dose	Number	Number
1	1	6	2
2	2	6	2
3	3	6	2
4	4	6	2

If all responses are uncorrelated with variance σ^2 then
Variance (dose $i -$ placebo) in cohort i is $\left(\frac{1}{6}+\frac{1}{2}\right)\sigma^2 = \frac{2}{3}\sigma^2$

From the protocol: "data of subjects having received placebo will be pooled in one group for analyses."

Variance (dose $i -$ placebo) is $\left(\frac{1}{6}+\frac{1}{8}\right)\sigma^2 = \frac{7}{24}\sigma^2$ if there are no cohort effects.

Variance (dose $i -$ dose j) is $\left(\frac{1}{6}+\frac{1}{6}\right)\sigma^2 = \frac{1}{3}\sigma^2$ if there are no cohort effects.

Are there cohort effects?

- Different types of people can volunteer at different times.

Are there cohort effects?

- Different types of people can volunteer at different times.
- There may be changes in the ambient conditions,
  eg temperature, pollutants, pollens.

Are there cohort effects?

- Different types of people can volunteer at different times.
- There may be changes in the ambient conditions,
  eg temperature, pollutants, pollens.
- The staff running the trial, or analysing the samples, may change.

Are there cohort effects?

- Different types of people can volunteer at different times.
- There may be changes in the ambient conditions,
  eg temperature, pollutants, pollens.
- The staff running the trial, or analysing the samples, may change.
- Protocols for using subsidiary equipment may change.

Are there cohort effects?

- Different types of people can volunteer at different times.
- There may be changes in the ambient conditions, eg temperature, pollutants, pollens.
- The staff running the trial, or analysing the samples, may change.
- Protocols for using subsidiary equipment may change.
- Halo effect among volunteers: if one reports nausea then they all may do so.

Are there cohort effects?

- Different types of people can volunteer at different times.
- There may be changes in the ambient conditions,
  eg temperature, pollutants, pollens.
- The staff running the trial, or analysing the samples, may change.
- Protocols for using subsidiary equipment may change.
- Halo effect among volunteers:
  if one reports nausea then they all may do so.
- Halo effect among staff:
  if they see symptoms in one volunteer, they expect them in others.

Are there cohort effects?

- Different types of people can volunteer at different times.
- There may be changes in the ambient conditions, eg temperature, pollutants, pollens.
- The staff running the trial, or analysing the samples, may change.
- Protocols for using subsidiary equipment may change.
- Halo effect among volunteers: if one reports nausea then they all may do so.
- Halo effect among staff: if they see symptoms in one volunteer, they expect them in others.

Are there cohort effects?

- Different types of people can volunteer at different times.
- There may be changes in the ambient conditions, eg temperature, pollutants, pollens.
- The staff running the trial, or analysing the samples, may change.
- Protocols for using subsidiary equipment may change.
- Halo effect among volunteers: if one reports nausea then they all may do so.
- Halo effect among staff: if they see symptoms in one volunteer, they expect them in others.

There have been many trials, in many topics, where, with hindsight, cohort effects swamp treatment effects.

Are there cohort effects?

- Different types of people can volunteer at different times.
- There may be changes in the ambient conditions, eg temperature, pollutants, pollens.
- The staff running the trial, or analysing the samples, may change.
- Protocols for using subsidiary equipment may change.
- Halo effect among volunteers: if one reports nausea then they all may do so.
- Halo effect among staff: if they see symptoms in one volunteer, they expect them in others.

There have been many trials, in many topics, where, with hindsight, cohort effects swamp treatment effects.
The Experimental Medicines Group of the Association of the British Pharmaceutical Industry (ABPI) says that trials should always be designed on the assumption that there will be cohort effects.

Analysis of the TeGenero trial with cohort effects

Cohort	TGN1412		Placebo
	Dose	Number	Number
1	1	6	2
2	2	6	2
3	3	6	2
4	4	6	2

$$\text{Variance (dose } i - \text{ placebo) in cohort } i = \left(\frac{1}{6} + \frac{1}{2}\right)\sigma^2 = \frac{2}{3}\sigma^2.$$

Analysis of the TeGenero trial with cohort effects

Cohort	TGN1412		Placebo
	Dose	Number	Number
1	1	6	2
2	2	6	2
3	3	6	2
4	4	6	2

$$\text{Variance (dose } i - \text{ placebo) in cohort } i = \left(\frac{1}{6} + \frac{1}{2}\right)\sigma^2 = \frac{2}{3}\sigma^2.$$

$$\begin{aligned}&\text{Estimator of (dose } i - \text{ dose } j) = \\ &\quad [\text{estimator of (dose } i - \text{ placebo) in cohort } i] - \\ &\quad [\text{estimator of (dose } j - \text{ placebo) in cohort } j]\end{aligned}$$

Analysis of the TeGenero trial with cohort effects

Cohort	TGN1412		Placebo
	Dose	Number	Number
1	1	6	2
2	2	6	2
3	3	6	2
4	4	6	2

$$\text{Variance (dose } i - \text{ placebo) in cohort } i = \left(\frac{1}{6} + \frac{1}{2}\right)\sigma^2 = \frac{2}{3}\sigma^2.$$

$$\begin{aligned}&\text{Estimator of (dose } i - \text{ dose } j) = \\ &\quad [\text{estimator of (dose } i - \text{ placebo) in cohort } i] - \\ &\quad [\text{estimator of (dose } j - \text{ placebo) in cohort } j]\end{aligned}$$

$$\text{So variance (dose } i - \text{ dose } j) = \left(\frac{2}{3} + \frac{2}{3}\right)\sigma^2 = \frac{4}{3}\sigma^2.$$

Senn's proposed design

Cohort	TGN1412		Placebo
	Dose	Number	Number
1	1	4	4
2	2	4	4
3	3	4	4
4	4	4	4

Senn's proposed design

Cohort	TGN1412		Placebo
	Dose	Number	Number
1	1	4	4
2	2	4	4
3	3	4	4
4	4	4	4

$$\text{Variance (dose } i - \text{ placebo) in cohort } i = \left(\frac{1}{4} + \frac{1}{4}\right)\sigma^2 = \frac{1}{2}\sigma^2$$

Senn's proposed design

Cohort	TGN1412		Placebo
	Dose	Number	Number
1	1	4	4
2	2	4	4
3	3	4	4
4	4	4	4

$$\text{Variance (dose } i - \text{ placebo) in cohort } i = \left(\frac{1}{4} + \frac{1}{4}\right)\sigma^2 = \frac{1}{2}\sigma^2 < \frac{2}{3}\sigma^2.$$

Senn's proposed design

Cohort	TGN1412		Placebo
	Dose	Number	Number
1	1	4	4
2	2	4	4
3	3	4	4
4	4	4	4

$$\text{Variance (dose } i - \text{ placebo) in cohort } i = \left(\frac{1}{4}+\frac{1}{4}\right)\sigma^2 = \frac{1}{2}\sigma^2 < \frac{2}{3}\sigma^2.$$

$$\text{So variance (dose } i - \text{ dose } j) = \left(\frac{1}{2}+\frac{1}{2}\right)\sigma^2 = \sigma^2$$

Senn's proposed design

Cohort	TGN1412		Placebo
	Dose	Number	Number
1	1	4	4
2	2	4	4
3	3	4	4
4	4	4	4

$$\text{Variance (dose } i - \text{ placebo) in cohort } i = \left(\frac{1}{4} + \frac{1}{4}\right)\sigma^2 = \frac{1}{2}\sigma^2 < \frac{2}{3}\sigma^2.$$

$$\text{So variance (dose } i - \text{ dose } j) = \left(\frac{1}{2} + \frac{1}{2}\right)\sigma^2 = \sigma^2 < \frac{4}{3}\sigma^2.$$

Senn's proposed design

Cohort	TGN1412		Placebo
	Dose	Number	Number
1	1	4	4
2	2	4	4
3	3	4	4
4	4	4	4

$$\text{Variance (dose } i - \text{ placebo) in cohort } i = \left(\frac{1}{4} + \frac{1}{4}\right)\sigma^2 = \frac{1}{2}\sigma^2 < \frac{2}{3}\sigma^2.$$

$$\text{So variance (dose } i - \text{ dose } j) = \left(\frac{1}{2} + \frac{1}{2}\right)\sigma^2 = \sigma^2 < \frac{4}{3}\sigma^2.$$

The TeGenero design is inadmissible because everything can be estimated, from the same resources, with smaller variance, by another design.

Dose-escalation trials: standard designs

There are n doses, with dose $1 <$ dose $2 < \cdots <$ dose n.

0 denotes the placebo.

There are n cohorts of m subjects each.

Dose-escalation trials: standard designs

There are n doses, with dose $1 <$ dose $2 < \cdots <$ dose n.

0 denotes the placebo.

There are n cohorts of m subjects each.

Dose-escalation trials: standard designs

There are n doses, with dose $1 <$ dose $2 < \cdots <$ dose n.

0 denotes the placebo.

There are n cohorts of m subjects each.

Cohort 1 subjects may receive only dose 1 or placebo.

In Cohort i, some subjects receive dose i;
no subject receives dose j if $j > i$.

Dose-escalation trials: standard designs

There are n doses, with dose $1 <$ dose $2 < \cdots <$ dose n.

0 denotes the placebo.

There are n cohorts of m subjects each.

Cohort 1 subjects may receive only dose 1 or placebo.

In Cohort i, some subjects receive dose i;
no subject receives dose j if $j > i$.

Put $s_{ki} =$ number of subjects who get dose i in cohort k. Then

$$s_{ki} > 0 \quad \text{if} \quad i = k$$
$$s_{ki} = 0 \quad \text{if} \quad i > k.$$

Scaled variance

Scaled variance

Assess designs by looking at the pairwise variances.

Scaled variance

Assess designs by looking at the pairwise variances.

If we double the number of subjects getting each dose in each cohort, then all variances are divided by 4. We want to know which pattern of design is good irrespective of the number of subjects.

Scaled variance

Assess designs by looking at the pairwise variances.

If we double the number of subjects getting each dose in each cohort, then all variances are divided by 4. We want to know which pattern of design is good irrespective of the number of subjects.

If doses could be equally replicated within each cohort, then each pairwise variance would be

$$\frac{2(n+1)\sigma^2}{\text{number of observations}}$$

Scaled variance

Assess designs by looking at the pairwise variances.

If we double the number of subjects getting each dose in each cohort, then all variances are divided by 4. We want to know which pattern of design is good irrespective of the number of subjects.

If doses could be equally replicated within each cohort, then each pairwise variance would be

$$\frac{2(n+1)\sigma^2}{\text{number of observations}}$$

so define the scaled variance v_{ij} to be

$$\frac{\text{Variance (dose } i - \text{ dose } j) \times \text{number of observations}}{2(n+1)\sigma^2}.$$

Textbook design

Aim:

- only doses 0 and k in cohort k
- equal replication overall.

Textbook design

Aim:

- only doses 0 and k in cohort k
- equal replication overall.

$$s_{ki} = \begin{cases} \dfrac{m}{n+1} & \text{if } i = 0 \\[2ex] \dfrac{nm}{n+1} & \text{if } 0 < i = k \\[2ex] 0 & \text{otherwise.} \end{cases}$$

Textbook design

Aim:

- only doses 0 and k in cohort k
- equal replication overall.

$$s_{ki} = \begin{cases} \dfrac{m}{n+1} & \text{if } i = 0 \\ \dfrac{nm}{n+1} & \text{if } 0 < i = k \\ 0 & \text{otherwise.} \end{cases}$$

Example: $n = 4$, $m = 10$

Dose	0	1	2	3	4
Cohort 1	2	8	0	0	0
Cohort 2	2	0	8	0	0
Cohort 3	2	0	0	8	0
Cohort 4	2	0	0	0	8

Textbook design

Aim:

- only doses 0 and k in cohort k
- equal replication overall.

$$s_{ki} = \begin{cases} \dfrac{m}{n+1} & \text{if } i = 0 \\ \dfrac{nm}{n+1} & \text{if } 0 < i = k \\ 0 & \text{otherwise.} \end{cases}$$

Example: $n = 4$, $m = 10$

Dose	0	1	2	3	4
Cohort 1	2	8	0	0	0
Cohort 2	2	0	8	0	0
Cohort 3	2	0	0	8	0
Cohort 4	2	0	0	0	8

$$v_{0i} = \frac{n+1}{2} \qquad v_{ij} = n+1$$

Senn's design

Aim:

- only doses 0 and k in cohort k
- minimize pairwise variances if there are cohort effects.

Senn's design

Aim:

- only doses 0 and k in cohort k
- minimize pairwise variances if there are cohort effects.

$$s_{ki} = \begin{cases} \frac{m}{2} & \text{if } i = 0 \\ \frac{m}{2} & \text{if } 0 < i = k \\ 0 & \text{otherwise.} \end{cases}$$

Senn's design

Aim:

- only doses 0 and k in cohort k
- minimize pairwise variances if there are cohort effects.

$$s_{ki} = \begin{cases} \frac{m}{2} & \text{if } i = 0 \\ \frac{m}{2} & \text{if } 0 < i = k \\ 0 & \text{otherwise.} \end{cases}$$

Example: $n = 4$, $m = 8$

Dose	0	1	2	3	4
Cohort 1	4	4	0	0	0
Cohort 2	4	0	4	0	0
Cohort 3	4	0	0	4	0
Cohort 4	4	0	0	0	4

Senn's design

Aim:

- only doses 0 and k in cohort k
- minimize pairwise variances if there are cohort effects.

$$s_{ki} = \begin{cases} \frac{m}{2} & \text{if } i = 0 \\ \frac{m}{2} & \text{if } 0 < i = k \\ 0 & \text{otherwise.} \end{cases}$$

Example: $n = 4$, $m = 8$

Dose	0	1	2	3	4
Cohort 1	4	4	0	0	0
Cohort 2	4	0	4	0	0
Cohort 3	4	0	0	4	0
Cohort 4	4	0	0	0	4

$$v_{0i} = \frac{2n}{n+1} \qquad v_{ij} = \frac{4n}{n+1}$$

Lessons from experience with block designs: I

The design is effectively a block design, with the cohorts as blocks.

Lessons from experience with block designs: I

The design is effectively a block design, with the cohorts as blocks.

Principle

In each cohort, no treatment should be allocated to more than half of the subjects.

Lessons from experience with block designs: I

The design is effectively a block design, with the cohorts as blocks.

Principle

In each cohort, no treatment should be allocated to more than half of the subjects.

Principle

Each cohort should have as many different treatments as possible.

Proposed "uniform halving" designs

Aim:

- make pairwise variances lower than in other designs, whether or not there are cohort effects.

Proposed “uniform halving” designs

Aim:

- make pairwise variances lower than in other designs, whether or not there are cohort effects.

$$s_{ki} = \begin{cases} \frac{m}{2} & \text{if } i = k \\ \text{nonzero} & \text{if } 0 \leq i < k \\ 0 & \text{otherwise.} \end{cases}$$

Proposed "uniform halving" designs

Aim:

- make pairwise variances lower than in other designs, whether or not there are cohort effects.

$$s_{ki} = \begin{cases} \frac{m}{2} & \text{if } i = k \\ \text{nonzero} & \text{if } 0 \le i < k \\ 0 & \text{otherwise.} \end{cases}$$

In Cohort 1: $\frac{m}{2}$ subjects get dose 1; $\frac{m}{2}$ subjects get placebo.

In Cohort k: $\frac{m}{2}$ subjects get dose k; remaining subjects are allocated as equally as possible to treatments 0 to $k-1$, with larger values given to make the 'replication so far' as equal as possible.

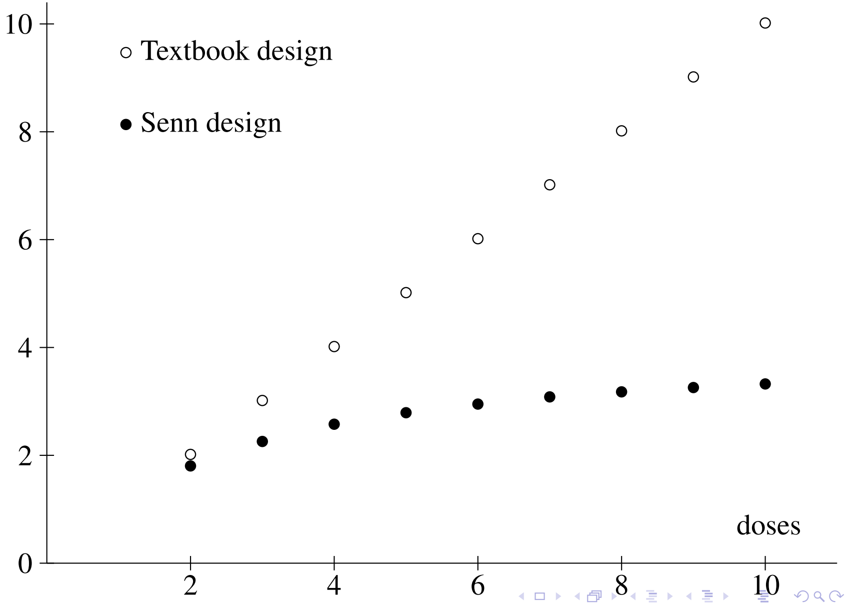
Example of a uniform halving design

Example: $n = 4$, $m = 8$

Dose	0	1	2	3	4
Cohort 1	4	4	0	0	0
Cohort 2	2	2	4	0	0
Cohort 3	1	1	2	4	0
Cohort 4	1	1	1	1	4

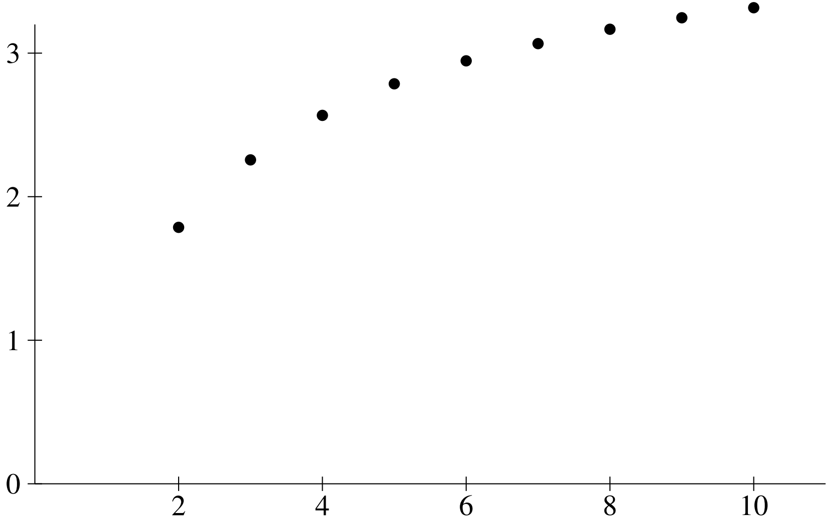
The scaled variances v_{ij} have to be calculated numerically.

Average scaled pairwise variance

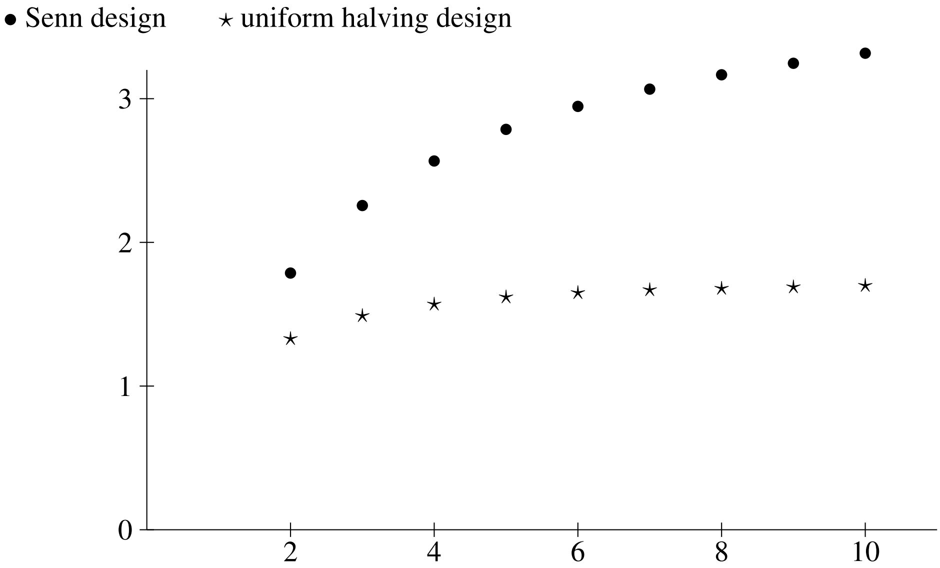


Average scaled pairwise variance: continued

- Senn design



Average scaled pairwise variance: continued



Lessons from experience with block designs: II

In the standard designs, the highest dose has all of its subjects in the final cohort.

Lessons from experience with block designs: II

In the standard designs, the highest dose has all of its subjects in the final cohort.

In ordinary block designs, you would never limit any treatment to just one block.

Principle

There should be one more cohort than there are doses, so that every dose can occur in at least two cohorts.

Dose-escalation trials: extended designs

There are n doses, with dose $1 <$ dose $2 < \cdots <$ dose n.

0 denotes the placebo.

There are $n+1$ cohorts of m subjects each.

Dose-escalation trials: extended designs

There are n doses, with dose $1 <$ dose $2 < \cdots <$ dose n.

0 denotes the placebo.

There are $n+1$ cohorts of m subjects each.

Dose-escalation trials: extended designs

There are n doses, with dose $1 <$ dose $2 < \cdots <$ dose n.

0 denotes the placebo.

There are $n+1$ cohorts of m subjects each.

Cohort 1 subjects may receive only dose 1 or placebo.

In Cohort i, for $2 \leq i \leq n$, some subjects receive dose i;
no subject receives dose j if $j > i$.

In Cohort $n+1$, any dose, or placebo, may be used.

Extended Senn design

In the final cohort,
compensate for the previous over-replication of placebo.

$$s_{n+1,i} = \begin{cases} 0 & \text{if } i = 0 \\ \dfrac{m}{n} & \text{otherwise} \end{cases}$$

Extended Senn design

In the final cohort,
compensate for the previous over-replication of placebo.

$$s_{n+1,i} = \begin{cases} 0 & \text{if } i = 0 \\ \dfrac{m}{n} & \text{otherwise} \end{cases}$$

Example: $n = 4$, $m = 8$

Dose	0	1	2	3	4
Cohort 1	4	4	0	0	0
Cohort 2	4	0	4	0	0
Cohort 3	4	0	0	4	0
Cohort 4	4	0	0	0	4
Cohort 5	0	2	2	2	2

Extended Senn design

In the final cohort,
compensate for the previous over-replication of placebo.

Example: $n = 4$, $m = 8$

$$s_{n+1,i} = \begin{cases} 0 & \text{if } i = 0 \\ \dfrac{m}{n} & \text{otherwise} \end{cases}$$

Dose	0	1	2	3	4
Cohort 1	4	4	0	0	0
Cohort 2	4	0	4	0	0
Cohort 3	4	0	0	4	0
Cohort 4	4	0	0	0	4
Cohort 5	0	2	2	2	2

$$v_{0i} = \frac{2(n^2+4)}{n(n+4)} \qquad v_{ij} = \frac{4n}{n+4}$$

Extension of the uniform halving design

About half the subjects in the final cohort are equally split between all treatments,
the remainder being allocated to make the overall replications as equal as possible, with any inequalities favouring the higher doses.

Extension of the uniform halving design

About half the subjects in the final cohort are equally split between all treatments,
the remainder being allocated to make the overall replications as equal as possible, with any inequalities favouring the higher doses.

Example: $n = 4$, $m = 8$

Dose	0	1	2	3	4
Cohort 1	4	4	0	0	0
Cohort 2	2	2	4	0	0
Cohort 3	1	1	2	4	0
Cohort 4	1	1	1	1	4

Extension of the uniform halving design

About half the subjects in the final cohort are equally split between all treatments,
the remainder being allocated to make the overall replications as equal as possible, with any inequalities favouring the higher doses.

Example: $n = 4$, $m = 8$

Dose	0	1	2	3	4
Cohort 1	4	4	0	0	0
Cohort 2	2	2	4	0	0
Cohort 3	1	1	2	4	0
Cohort 4	1	1	1	1	4
	1	1	1	1	1

Extension of the uniform halving design

About half the subjects in the final cohort are equally split between all treatments,
the remainder being allocated to make the overall replications as equal as possible, with any inequalities favouring the higher doses.

Example: $n = 4$, $m = 8$

Dose	0	1	2	3	4
Cohort 1	4	4	0	0	0
Cohort 2	2	2	4	0	0
Cohort 3	1	1	2	4	0
Cohort 4	1	1	1	1	4
	1	1	1	1	1
					1

Extension of the uniform halving design

About half the subjects in the final cohort are equally split between all treatments,
the remainder being allocated to make the overall replications as equal as possible, with any inequalities favouring the higher doses.

Example: $n = 4$, $m = 8$

Dose	0	1	2	3	4
Cohort 1	4	4	0	0	0
Cohort 2	2	2	4	0	0
Cohort 3	1	1	2	4	0
Cohort 4	1	1	1	1	4
	1	1	1	1	1
					1
				1	1

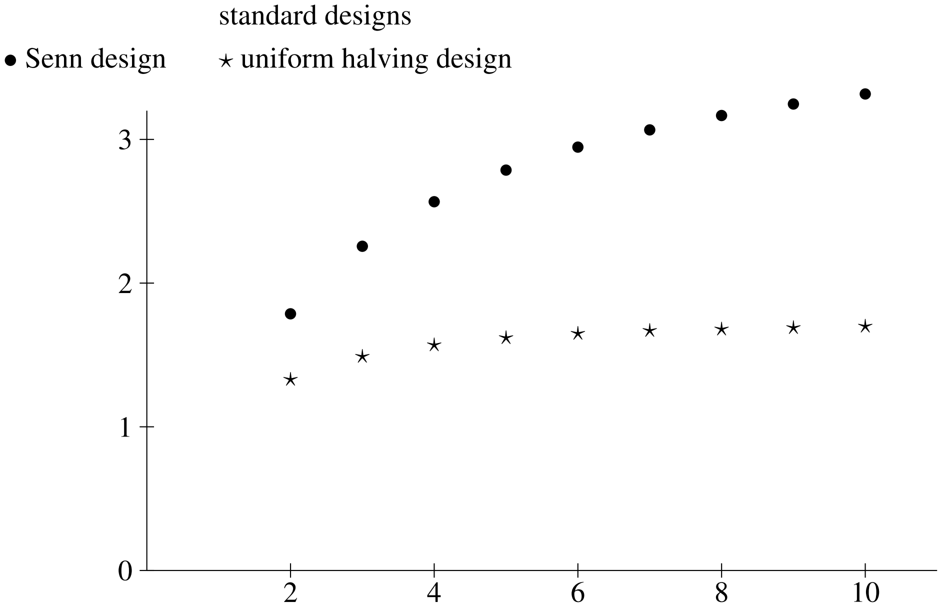
Extension of the uniform halving design

About half the subjects in the final cohort are equally split between all treatments,
the remainder being allocated to make the overall replications as equal as possible, with any inequalities favouring the higher doses.

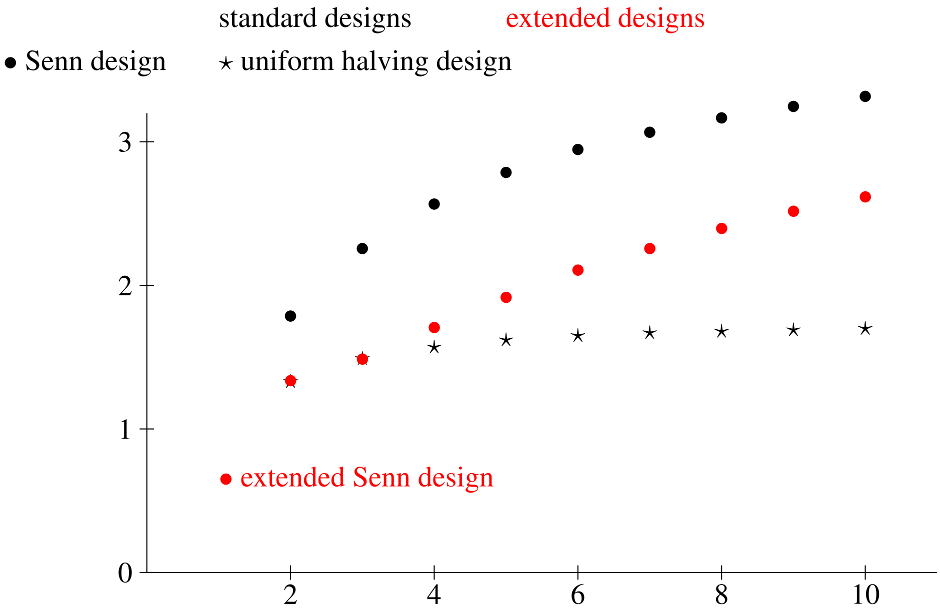
Example: $n = 4$, $m = 8$

Dose	0	1	2	3	4
Cohort 1	4	4	0	0	0
Cohort 2	2	2	4	0	0
Cohort 3	1	1	2	4	0
Cohort 4	1	1	1	1	4
	1	1	1	1	1
					1
				1	1
Cohort 5	1	1	1	2	3

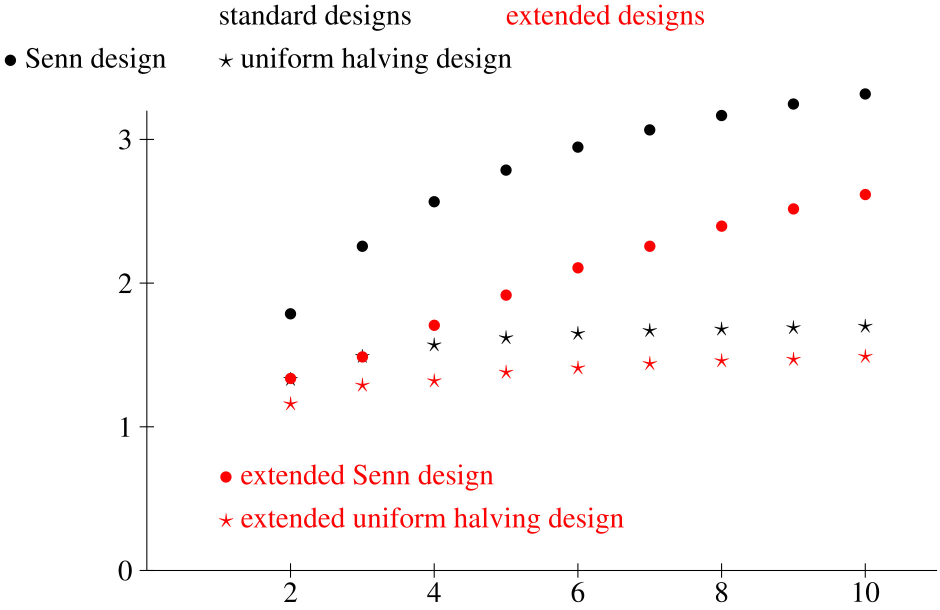
Average scaled pairwise variance: continued (again)



Average scaled pairwise variance: continued (again)



Average scaled pairwise variance: continued (again)



Two designs for 4 doses using 40 subjects

Numbers of subjects

Std TB

Dose	0	1	2	3	4
Cohort 1	2	8	0	0	0
Cohort 2	2	0	8	0	0
Cohort 3	2	0	0	8	0
Cohort 4	2	0	0	0	8

Actual pairwise variances/σ^2

	1	2	3	4
0	0.625	0.625	0.625	0.625
1		1.250	1.250	1.250
2			1.250	1.250
3				1.250

Ext UH

Dose	0	1	2	3	4
Cohort 1	4	4	0	0	0
Cohort 2	2	2	4	0	0
Cohort 3	1	1	2	4	0
Cohort 4	1	1	1	1	4
Cohort 5	1	1	1	2	3

	1	2	3	4
0	0.222	0.285	0.348	0.370
1		0.285	0.348	0.370
2			0.330	0.378
3				0.375

Two designs for 4 doses using 40 subjects

Std TB

Numbers of subjects

Dose	0	1	2	3	4
Cohort 1	2	8	0	0	0
Cohort 2	2	0	8	0	0
Cohort 3	2	0	0	8	0
Cohort 4	2	0	0	0	8

Actual pairwise variances/σ^2

	1	2	3	4
0	0.625	0.625	0.625	0.625
1		1.250	1.250	1.250
2			1.250	1.250
3				1.250

average 1.00

Ext UH

Dose	0	1	2	3	4
Cohort 1	4	4	0	0	0
Cohort 2	2	2	4	0	0
Cohort 3	1	1	2	4	0
Cohort 4	1	1	1	1	4
Cohort 5	1	1	1	2	3

	1	2	3	4
0	0.222	0.285	0.348	0.370
1		0.285	0.348	0.370
2			0.330	0.378
3				0.375

average 0.33

Simple rule

Among the standard designs examined, the uniform halving designs are best.

Simple rule

Among the standard designs examined, the uniform halving designs are best.

Among the extended designs examined, the best are the uniform halving designs with the particular extension given.

Simple rule

Among the standard designs examined, the uniform halving designs are best.

Among the extended designs examined, the best are the uniform halving designs with the particular extension given.

Both types can be described by the following simple rule:

Simple rule

Among the standard designs examined, the uniform halving designs are best.

Among the extended designs examined, the best are the uniform halving designs with the particular extension given.

Both types can be described by the following simple rule:

Principle

In each cohort,
half of the subjects should be distributed (approximately) equally among all the treatments that have been used in any previous cohort;
the remaining subjects should be used to make the replication so far as equal as possible by compensating for previous under-replication.

Advantages of the halving designs

- Variance is reduced by a factor of two or more.

Advantages of the halving designs

- Variance is reduced by a factor of two or more.
- The allocation rule is simple,
  and can be applied to any number of subjects per cohort.

Advantages of the halving designs

- Variance is reduced by a factor of two or more.
- The allocation rule is simple,
  and can be applied to any number of subjects per cohort.
- If the trial has to be stopped early because dose i is harmful,
  then fewer subjects will have been exposed to dose i than would have been with the textbook design.

Advantages of the halving designs

- Variance is reduced by a factor of two or more.
- The allocation rule is simple,
  and can be applied to any number of subjects per cohort.
- If the trial has to be stopped early because dose i is harmful, then fewer subjects will have been exposed to dose i than would have been with the textbook design.
- If the trial has to be stopped early because dose i is harmful, then the previous $i-1$ cohorts form the recommended standard design for $i-1$ doses; if desired, they can be followed by an extra cohort for treatments $0, \ldots, i-1$ only.

Advantages of the halving designs

- Variance is reduced by a factor of two or more.
- The allocation rule is simple,
  and can be applied to any number of subjects per cohort.
- If the trial has to be stopped early because dose i is harmful, then fewer subjects will have been exposed to dose i than would have been with the textbook design.
- If the trial has to be stopped early because dose i is harmful, then the previous $i-1$ cohorts form the recommended standard design for $i-1$ doses; if desired, they can be followed by an extra cohort for treatments $0, \ldots, i-1$ only.
- If cohort effects are small and random, the variance is very little more than for the textbook design (not shown here).

Advantages of the halving designs

- Variance is reduced by a factor of two or more.
- The allocation rule is simple,
  and can be applied to any number of subjects per cohort.
- If the trial has to be stopped early because dose i is harmful, then fewer subjects will have been exposed to dose i than would have been with the textbook design.
- If the trial has to be stopped early because dose i is harmful, then the previous $i-1$ cohorts form the recommended standard design for $i-1$ doses; if desired, they can be followed by an extra cohort for treatments $0, \ldots, i-1$ only.
- If cohort effects are small and random, the variance is very little more than for the textbook design (not shown here).
- Blinding is more effective than in textbook designs.

Conclusion

- Remove bias by

Conclusion

- Remove bias by
  - identifying suitable blocks and using them in the design and in the analysis;

Conclusion

- Remove bias by
  - identifying suitable blocks and using them in the design and in the analysis;
  - randomizing appropriately to remove unknown sources of bias.

Conclusion

- Remove bias by
  - identifying suitable blocks and using them in the design and in the analysis;
  - randomizing appropriately to remove unknown sources of bias.
- Reduce variance by choosing the best combination of design and replications.

Conclusion

- Remove bias by
  - identifying suitable blocks and using them in the design and in the analysis;
  - randomizing appropriately to remove unknown sources of bias.
- Reduce variance by choosing the best combination of design and replications.
- Don't be afraid to transfer design principles from one area of science to another.