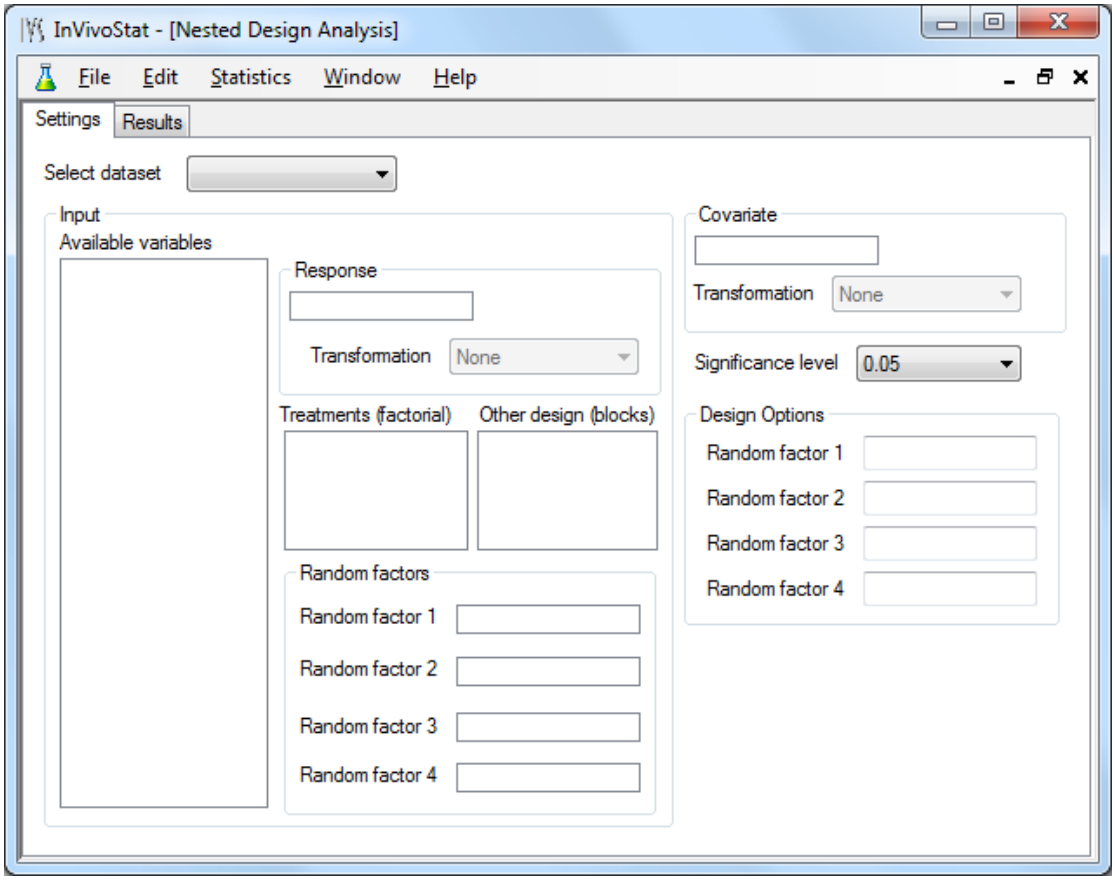
InVivoStat

Nested Design Analysis Module

Tipsheet

The Nested Design Analysis module in InVivoStat is available from the Unvalidated Analyses submenu in the Statistics drop-down menu entitled 'Nested Design Analysis'.



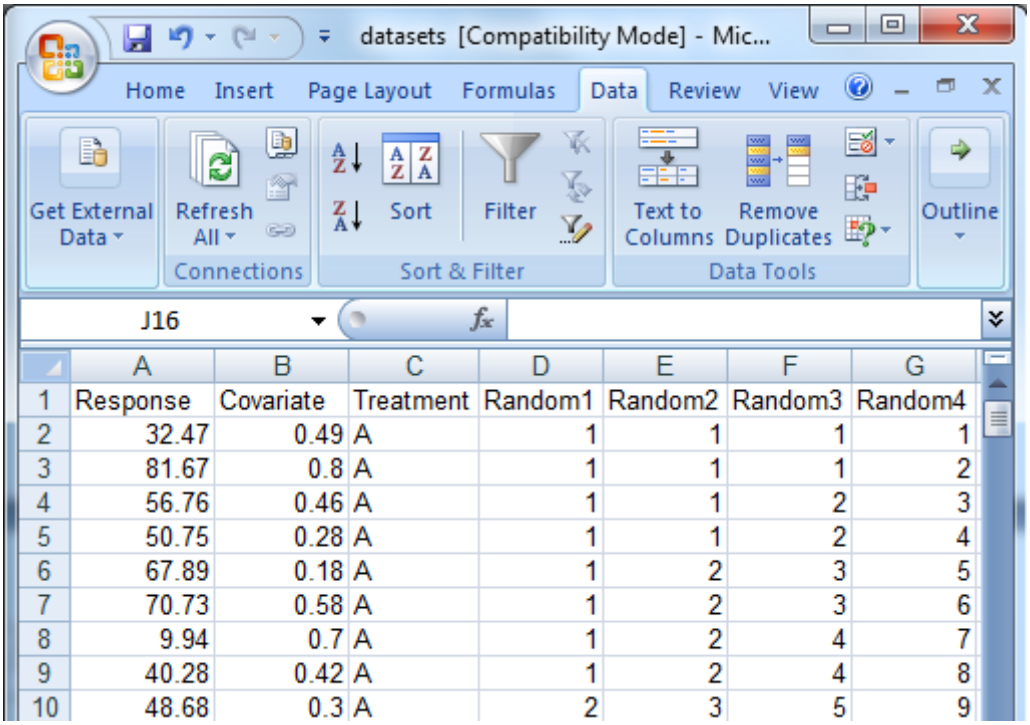
InVivoStat's Nested Design Analysis module is a tool for assessing the size and effect of the random effects. It can be employed when the researcher has used a hierarchical nested design (see Bate and Clark 2014, Section 3.7.3).

This module assumes the user wishes to compare treatments against the highest source of variability (Random factor 1). It allows the user to investigate how varying the replication of the random effects influences the statistical power of the treatment comparisons.

In other words, if the researcher increases the replication of the within-animal measurements, can they reduce the total number of animals used without decreasing the statistical power?
The analysis performed is described in Bate and Clark (2014, Section 3.7.3 and Section 6.14)

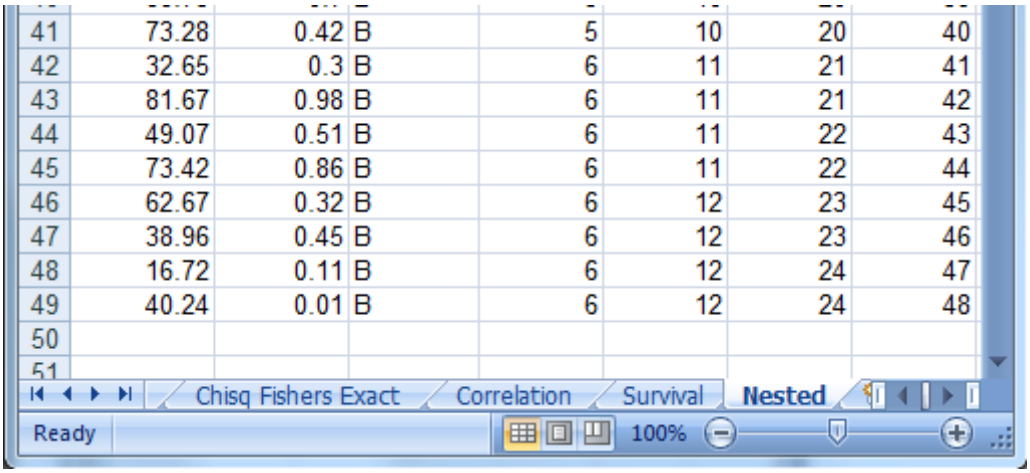
1 Setting up the dataset

This module required the dataset, and the variables within the dataset, to be in a specific format. It is recommended that this be completed before importing the data into InVivoStat. An example of the dataset, consisting of a response, treatment factor, a covariate and four random factors is given below.



	A	B	C	D	E	F	G
1	Response	Covariate	Treatment	Random1	Random2	Random3	Random4
2	32.47	0.49	A	1	1	1	1
3	81.67	0.8	A	1	1	1	2
4	56.76	0.46	A	1	1	2	3
5	50.75	0.28	A	1	1	2	4
6	67.89	0.18	A	1	2	3	5
7	70.73	0.58	A	1	2	3	6
8	9.94	0.7	A	1	2	4	7
9	40.28	0.42	A	1	2	4	8
10	48.68	0.3	A	2	3	5	9

• • •



41	73.28	0.42	B	5	10	20	40
42	32.65	0.3	B	6	11	21	41
43	81.67	0.98	B	6	11	21	42
44	49.07	0.51	B	6	11	22	43
45	73.42	0.86	B	6	11	22	44
46	62.67	0.32	B	6	12	23	45
47	38.96	0.45	B	6	12	23	46
48	16.72	0.11	B	6	12	24	47
49	40.24	0.01	B	6	12	24	48
50							
51							

The experiment consists of two treatments (factor Treatment: levels A and B), with three animals per treatment (factor Random 1: levels 1 - 6). There is then:

- Random 2 (nested within Random 1) consisting of 12 levels
- Random 3 (nested within Random 2) consisting of 24 levels
- Random 4 (nested within Random 3) consisting of 48 levels

Note that the levels are uniquely labelled (so the animals are numbered 1 to 6 and not 1 to 3 for Treatment A and 1 to 3 for treatment B). This is important.

Also note:

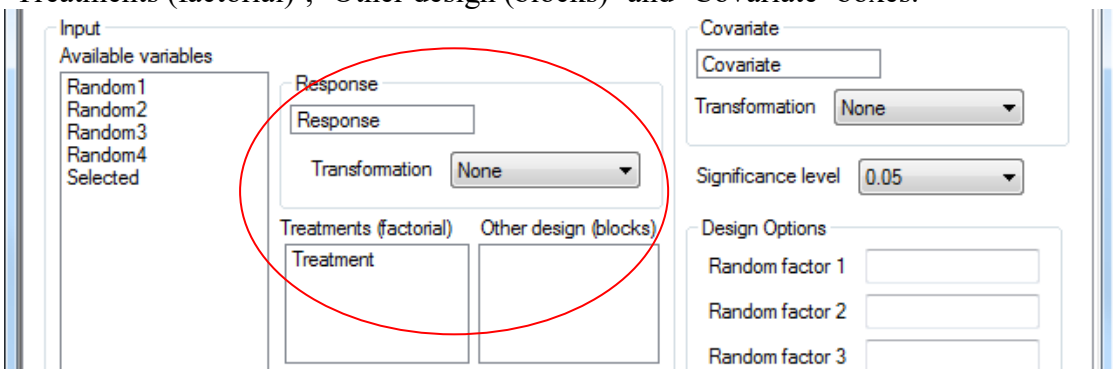
- All of the responses are contained within a single column, i.e. multiple rows per animal (similar to repeated measures datasets).
- You can include multiple treatment factors, blocking factors and a covariate, similar to other parametric analysis modules.
- There needs to be a random effect (Random 4 in this case) whose levels correspond to the individual observations.

2 Analysis procedure

Using InVivoStat analysis involves a series of stages:

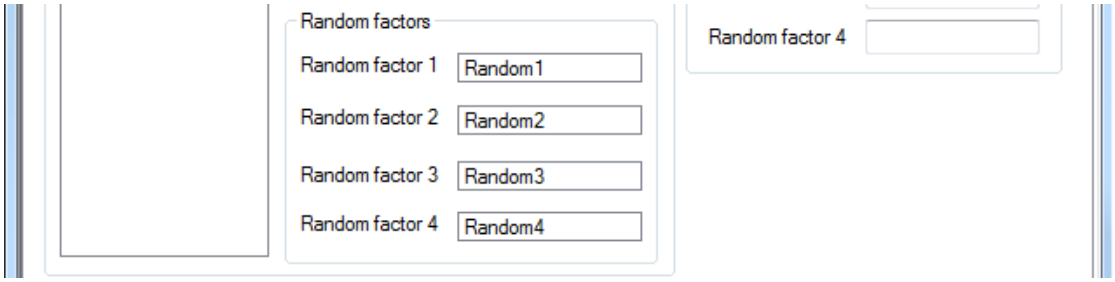
2.1 Setting up the model

Once the dataset has been opened, the user can select the variables for the analysis by dragging and dropping them from the 'Available variables' list into the 'Response', 'Treatments (factorial)', 'Other design (blocks)' and 'Covariate' boxes.



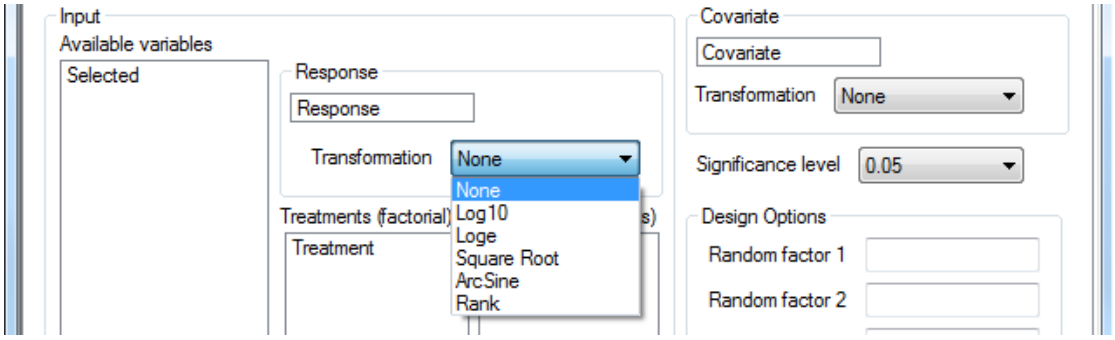
The user also selects the hierarchically nested random factors by dragging and dropping them into the relevant boxes. Note the order they are included is important:

- Random factor 2 is nested within Random factor 1,
- Random factor 3 is nested within Random factor 2,
- Random factor 4 is nested within Random factor 3.



2.2 Transforming the response/covariate

The user can transform the response/covariate if they believe the responses are not normally distributed.

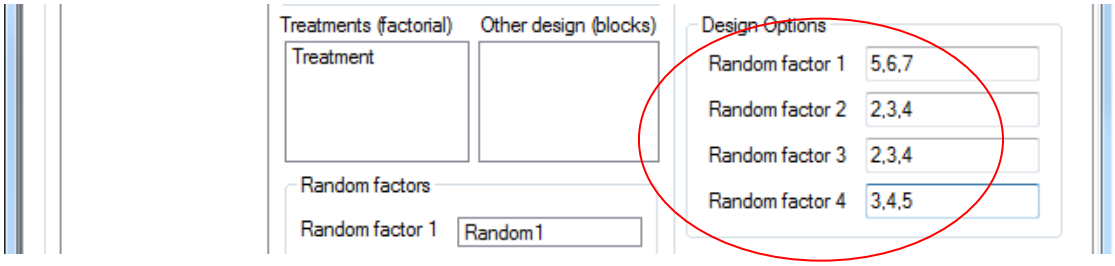


2.3 Selecting the significance level

The user can select the significance level they which to use in the power analysis, the default is 0.05 or 5%.

2.4 Selecting the design options to be investigated

Finally the user should select the levels of replication of the random factors to be investigated. Different levels of replication are entered, separated by commas.



3 Output results

The output consists of a series of plots (one per random factor in the nested design) that shows the increase in power that can be gained by increasing the replication of that random effect (while keeping other random effect replication fixed). Note the power plots produced are not orientated in the same way as the Power Analysis module plots. In this module the X-axis corresponds to the size of the difference whereas in the Power Analysis module the X-axis corresponds to the sample size.

In particular the output consists of:

Response and Covariate
InVivoStat identifies the response being analysed and also the covariate (if one is selected). This section also describes any transformations that have been applied.

Methodology
A description of the methodology applied is given.

Table of estimated variance components
This table summarises the breakdown of the total variability into parts corresponding to the nested random factors.

Table of average replication in the original design
This table describes the replication of the factors that will be used in the subsequent power analysis. When the replication of one of the factors is varied, the others are held constant at the level given in this table.

Power curve for an approximate original design
Given the replication levels in the "Table of average replication", a plot containing a single power curve is generated to illustrate the power of a design that consists of these replication levels.

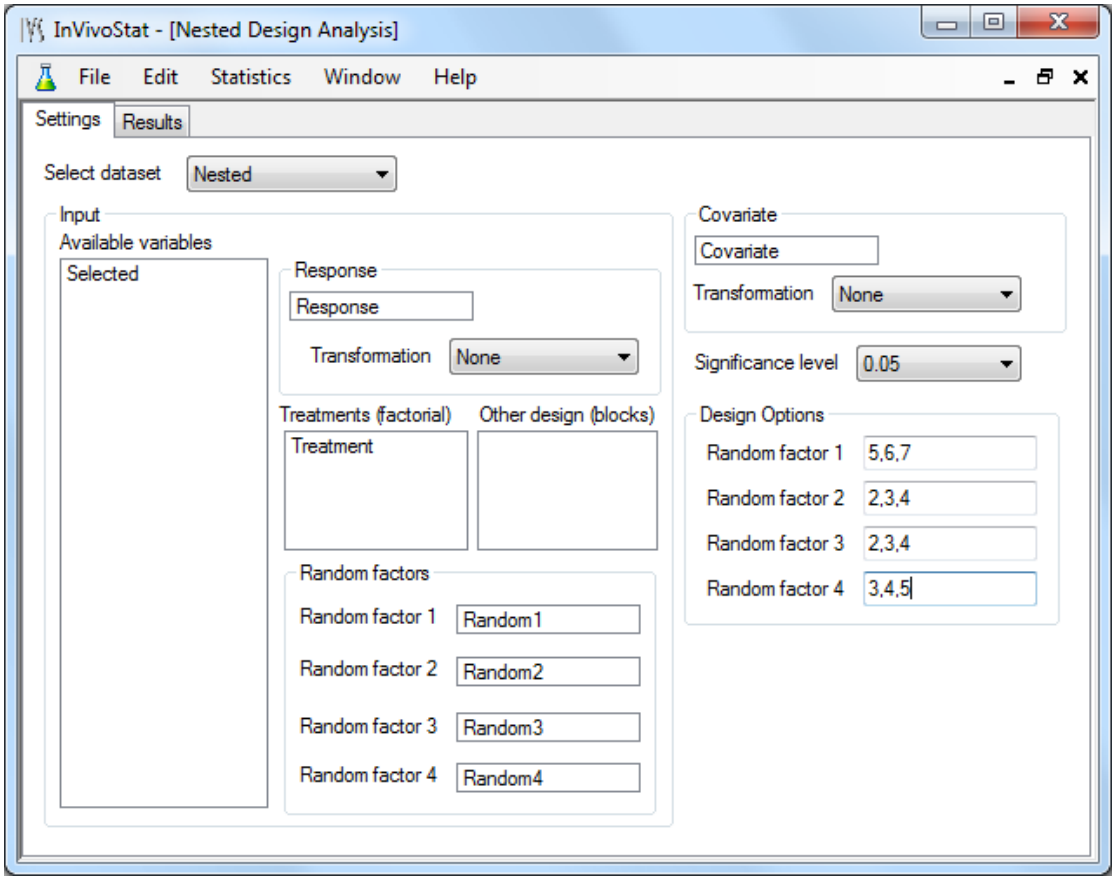
Power curves for a design varying the replication of Random *x*
For each of the nested random factors in the nested design, a separate plot is generated which consists of a series of power curves, one per level of replication of the given random factor (as defined by the user). The other random factors are held fixed, as described in the "Table of average replication".

Combined power curved plot
If all the user defined lists of replication levels, as defined on the interface, are of the same length, then InVivoStat produces a final plot that corresponds to varying all the random effects (using the levels given by the user).

Statistical and R references
Finally a list of relevant references is given.

4 Sample output

Options:



InVivoStat Nested Design Analysis

Warning

Warning: This module is currently under construction, care should be taken when considering the results. The results have not been verified.

Response and covariate

The Response response is currently being analysed by the Nested Design Analysis module, with "Covariate" fitted as a covariate.

Methodology

This module uses the estimated variance components from the original dataset to predict the hypothetical statistical power than can be achieved by varying the replication of the levels of the random factors in the experimental design, see Snedecor and Cochran (1989, p239).

For more information on the theoretical approaches that are implemented within this module, see Bate and Clark (2014).

Table of estimated variance components

	Random1	Random2	Random3	Random4
Variance Components	58.3860	0.1404	0.0042	599.7047

Table of average replication in the original design

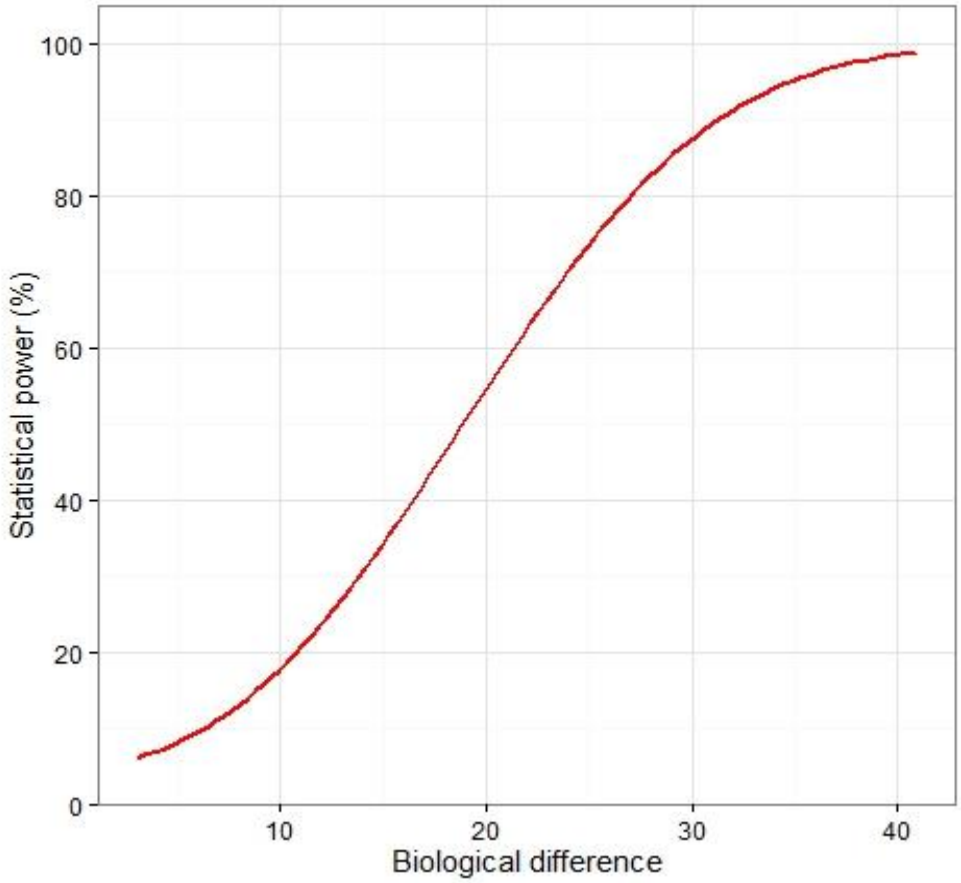
The following replication of the levels of the random factors are used in the power analyses, unless alternative replications are defined by the user. They are an estimate of the replication used within the original design.

	Random1	Random2	Random3	Random4
Factor replication	3	2	2	2

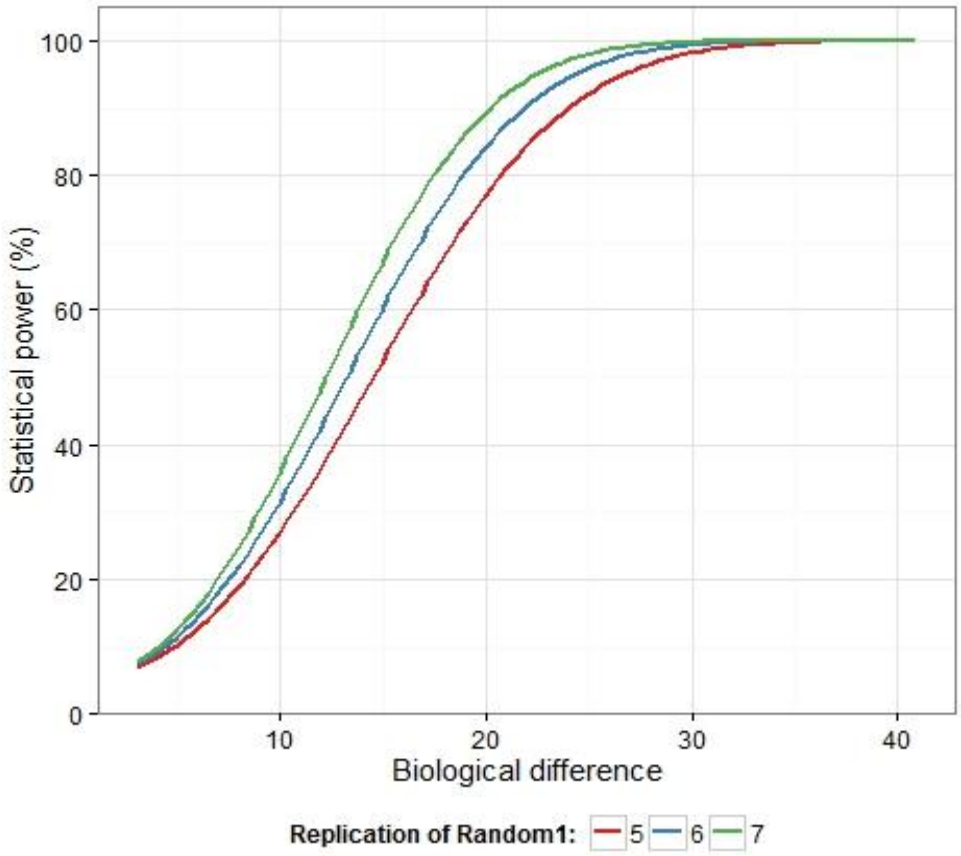
Estimated between subject variance

The estimate of the between subject variance is 1067.36. This is computed using the variance components, calculated from the data generated by the original design, and the average replication of the original design.

Power curve for an approximate original design

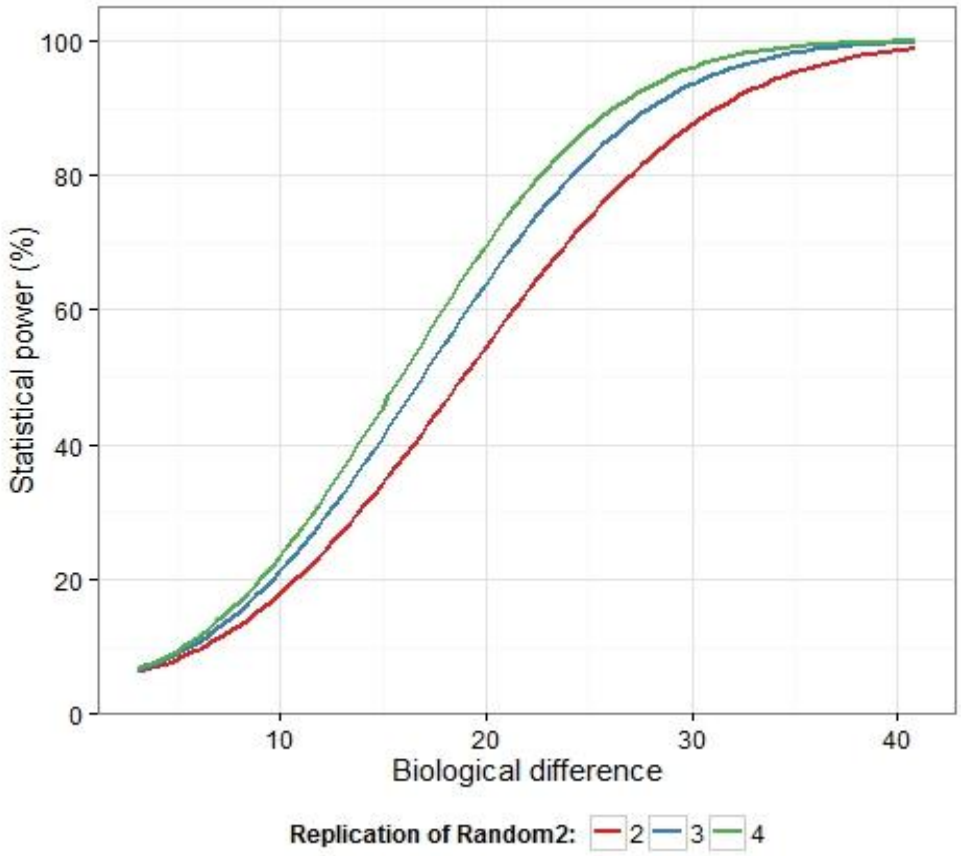


Power curves for a design varying the replication of Random1



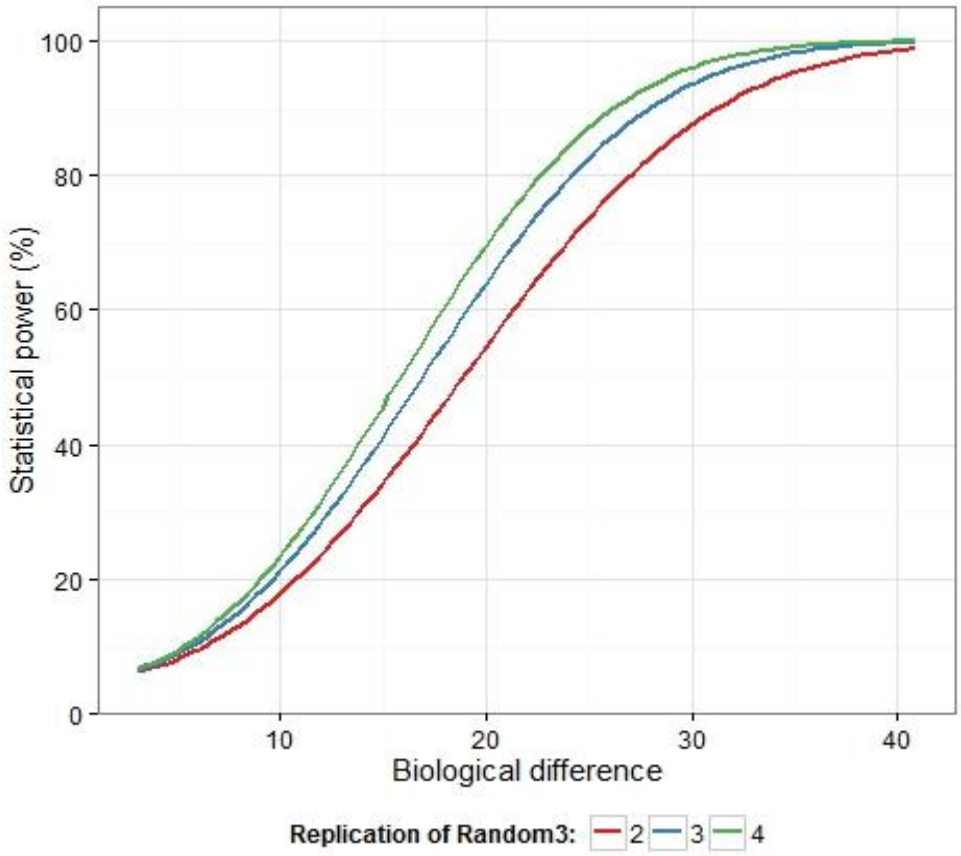
These power curves assume the replication of the levels of the random factor Random2 is 2, the replication of the levels of the random factor Random3 is 2 and the replication of the levels of the random factor Random4 is 2.

Power curves for a design varying the replication of Random2 within Random1



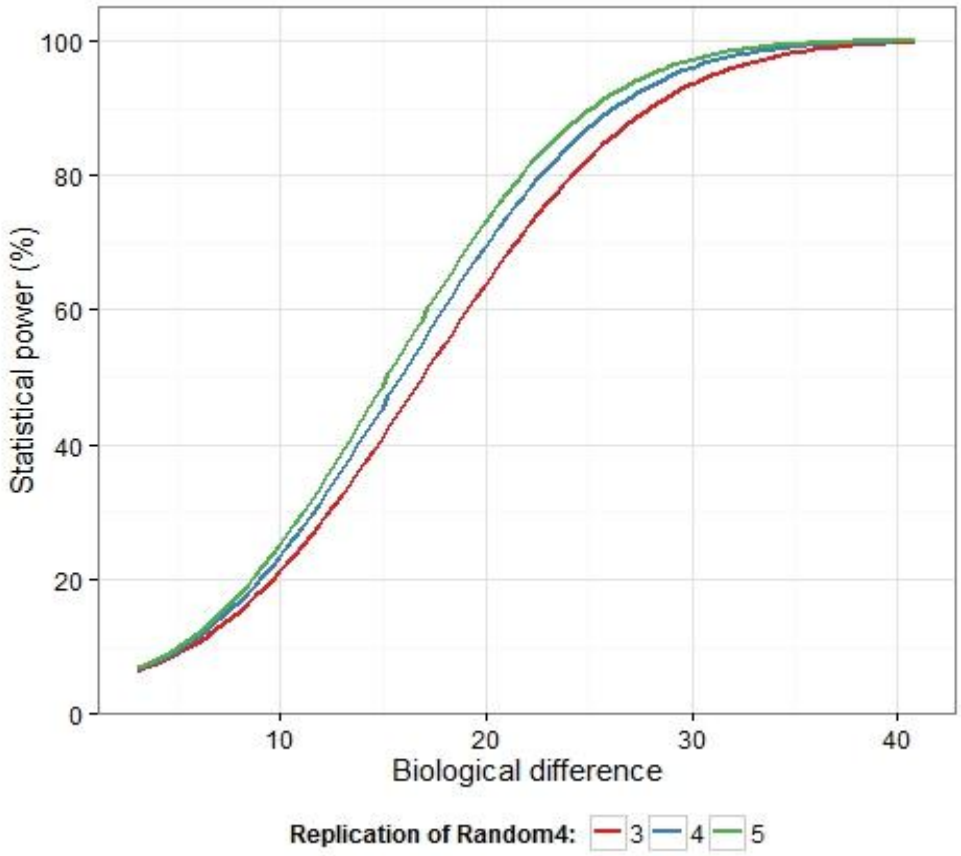
These power curves assume the replication of the levels of the random factor Random1 is 3, the replication of the levels of the random factor Random3 is 2 and the replication of Random4 is 2.

Power curves for a design varying the replication of Random3 within Random2 within Random1



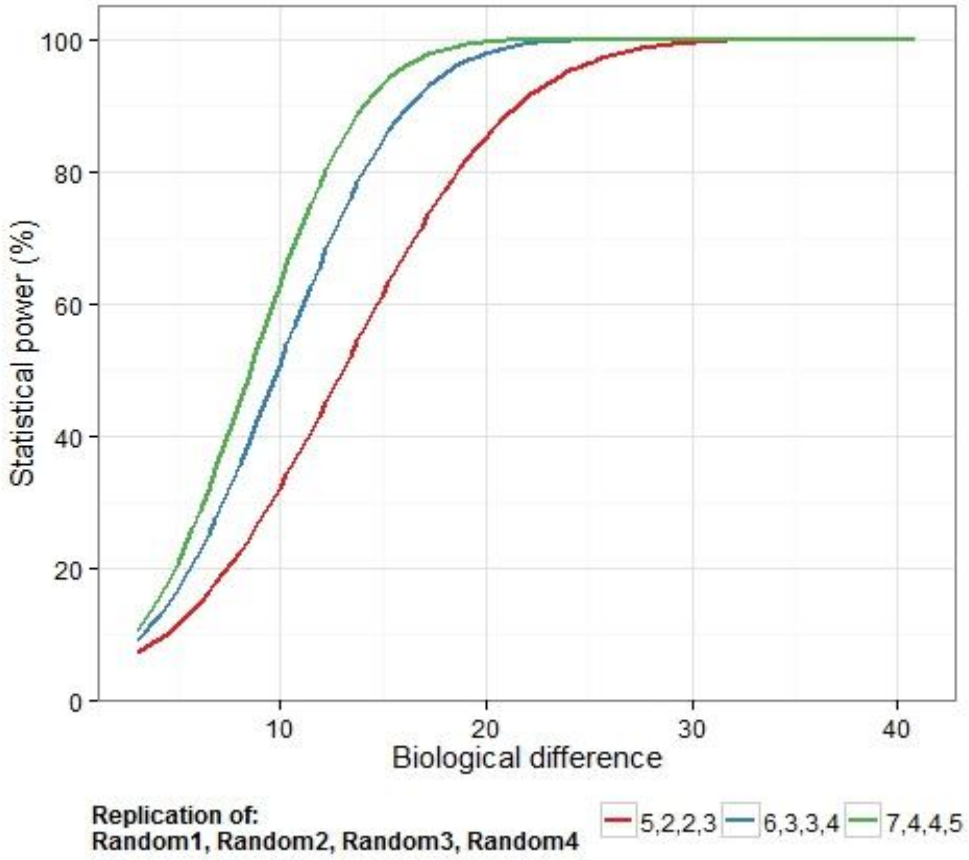
These power curves assume the replication of the levels of the random factor Random1 is 3, the replication of the levels of the random factor Random2 is 2 and the replication of Random4 is 2.

Power curves for a design varying the replication of Random4 within Random3 within Random2 within Random1



These power curves assume the replication of the levels of the random factor Random1 is 3, the replication of the levels of the random factor Random2 is 2 and the replication of Random3 is 2.

Power curves for a design varying the replication of Random1, Random2 within Random1, Random3 within Random2 within Random1 and Random4 within Random3 within Random2 within Random1



Statistical references

Bate ST and Clark RA. (2014). The Design and Statistical Analysis of Animal Experiments. Cambridge University Press.

Snedecor GW and Cochran WG. (1989). Statistical Methods (8th Edition). Iowa State University Press, Ames, Iowa.

R references

R Development Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org.

Barret Schloerke, Jason Crowley, Di Cook, Heike Hofmann, Hadley Wickham, Francois Briatte, Moritz Marbach and Edwin Thoen (2014). GGally: Extension to ggplot2. R package version 0.4.5. http://CRAN.R-project.org/package=GGally

Erich Neuwirth (2011). RColorBrewer: ColorBrewer palettes. R package version 1.0-5. http://CRAN.R-project.org/package=RColorBrewer

H. Wickham. ggplot2: elegant graphics for data analysis. Springer New York, 2009.

H. Wickham. Reshaping data with the reshape package. Journal of Statistical Software, 21(12), 2007.

Hadley Wickham (2011). The Split-Apply-Combine Strategy for Data Analysis. Journal of Statistical Software, 40(1), 1-29. URL http://www.jstatsoft.org/v40/i01/.

Hadley Wickham (2012). scales: Scale functions for graphics. R package version 0.2.3. http://CRAN.R-project.org/package=scales

Jose Pinheiro, Douglas Bates, Saikat DebRoy, Deepayan Sarkar and the R Development Core Team (2013). nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-111.

Lecoutre, Eric (2003). The R2HTML Package. R News, Vol 3. N. 3, Vienna, Austria.

Louis Kates and Thomas Petzoldt (2012). proto: Prototype object-based programming. R package version 0.3-10. http://CRAN.R-project.org/package=proto

Analysis dataset

	Response	Covariate	Treatment	Random1	Random2	Random3	Random4
1	32.5	0.49	A	1	1	1	1
2	81.7	0.80	A	1	1	1	2
3	56.8	0.46	A	1	1	2	3
4	50.8	0.28	A	1	1	2	4
5	67.9	0.18	A	1	2	3	5
6	70.7	0.58	A	1	2	3	6
7	9.9	0.70	A	1	2	4	7
8	40.3	0.42	A	1	2	4	8
9	48.7	0.30	A	2	3	5	9
10	46.0	0.98	A	2	3	5	10
11	47.4	0.51	A	2	3	6	11
12	22.4	0.86	A	2	3	6	12
13	4.6	0.32	A	2	4	7	13
14	9.5	0.45	A	2	4	7	14
15	16.4	0.11	A	2	4	8	15
16	12.2	0.01	A	2	4	8	16
17	28.9	0.49	A	3	5	9	17
18	93.2	0.80	A	3	5	9	18
19	42.6	0.46	A	3	5	10	19
20	6.3	0.28	A	3	5	10	20
21	70.7	0.18	A	3	6	11	21
22	21.6	0.58	A	3	6	11	22
23	99.9	0.70	A	3	6	12	23
24	69.6	0.42	A	3	6	12	24
25	95.6	0.30	B	4	7	13	25
26	32.6	0.98	B	4	7	13	26
27	99.4	0.51	B	4	7	14	27
28	64.1	0.86	B	4	7	14	28
29	35.6	0.32	B	4	8	15	29
30	71.0	0.45	B	4	8	15	30
31	43.8	0.11	B	4	8	16	31

32	61.6	0.01	B	4	8	16	32
33	7.6	0.49	B	5	9	17	33
34	80.5	0.80	B	5	9	17	34
35	41.6	0.46	B	5	9	18	35
36	59.0	0.28	B	5	9	18	36
37	65.8	0.18	B	5	10	19	37
38	57.6	0.58	B	5	10	19	38
39	88.8	0.70	B	5	10	20	39
40	73.3	0.42	B	5	10	20	40
41	32.6	0.30	B	6	11	21	41
42	81.7	0.98	B	6	11	21	42
43	49.1	0.51	B	6	11	22	43
44	73.4	0.86	B	6	11	22	44
45	62.7	0.32	B	6	12	23	45
46	39.0	0.45	B	6	12	23	46
47	16.7	0.11	B	6	12	24	47
48	40.2	0.01	B	6	12	24	48