CLUSTER SAMPLE SIZE CALCULATOR

USER MANUAL

Health Services Research Unit
University of Aberdeen
Polwarth Building
Foresterhill
ABERDEEN
AB25 2ZD
UK

Tel: +44 (0)1224 663123 extn 53909

May 1999

CLUSTER SAMPLE SIZE CALCULATOR GUIDE

Introduction

Cluster randomised trials involve randomisation of groups of individuals such as randomisation by general practice, hospital ward or health professional.

A fundamental assumption of the patient-randomised trial is that the outcome for an individual patient is completely unrelated to that for any other patient – they are said to be 'independent'. This assumption is violated when cluster randomisation is adopted, because patients within any one cluster are more likely to respond in a similar manner. For example, the management of patients in a single hospital is more likely to be consistent than management across a number of hospitals.

A statistical measure of this intracluster dependence is the 'intracluster correlation coefficient' (ICC) which is based on the relationship of the between to within-cluster variance. For example, in a study which randomised by hospital, the ICC would be high if the management of patients within hospitals was very consistent but there was wide variation in practice across hospitals.

Standard sample size formulae also assume that the outcomes for each patient are independent. With cluster RCTs, the use of these formulae will result in sample size estimates which will be too small, resulting in under-powered studies.

The Health Services Research Unit at Aberdeen University has developed software to help address this problem. Sample Size Calculator (SSC) is a Windows based software package that will make corrections to an unadjusted sample size. Accounting for ICC and cluster size, for both continuous and binary data, SSC will give the number of clusters of a certain size needed to detect a significance difference between to equally sized groups .

Sample Size Calculations

To start a calculation, click on File and select New Sample Size. This will bring up the following window.

Means

Say we wish to test for a difference between two groups of continuous data, control and experiment, for a certain intervention. The intervention, a clinical guideline written for the management of blood pressure, is to be randomised at the ward level, and all relevant patients in the active ward managed according to the guideline. A drop of 5mm Hg is deemed to be the clinically significant difference, previous literature estimates the standard deviation of blood pressure in each group is estimated at 15 mm Hg . We wish to know the required sample size to detect this difference (with 5% significance and 80% power). We know that the average cluster (ward) size is 15.

First of all click on the Means tab to select a calculation for continuous data. These values are entered into the sample size calculator by clicking in the appropriate box and typing in the values:

- 5 in *Min.Diff* (minimum difference detectable)
- 15 in *Std.Dev* (standard deviation)

The power and significance default settings are 5% and 80% respectively.

Click on *Calculate* to carry out the calculation.

The *Unadjusted* box now gives the total sample size required to detect the difference, 282 patients, ignoring the effect of clustering.

The table contains the *total* number of clusters (assuming a two-arm trial) needed for differing ICCs and cluster sizes. Cluster sizes are along the top and ICCs are listed down the side. The cluster size is the average cluster size, in this case average ward size. The ICC needs to be estimated, either from literature, from a pilot study or from similar studies in the past. At present there is little empirical evidence of sizes and factors affecting size, although there is a database of ICC estimates calculated from implementation datasets available through the HSRU web-site address: **http://www.abdn.ac.uk/public_health/hsru/sampsize.html.**

Let us assume an ICC of 0.01. From the table we can see that for an ICC of 0.01, and a cluster size of 15, the number of clusters required to achieve the adequate statistical power is 22 (resulting in an increase in the number of patients required from 282 to 330).

The sample size calculator uses the Design Effect [1] or Variance Inflation Factor [2] formula to make adjustments to the standard sample size calculations (see appendix for details).

The cluster sizes need not be confined to the pre-set values. Clicking on the *pencil* icon brings up small boxes above the columns. Entering values here allows us to define cluster sizes that differ from the default settings.
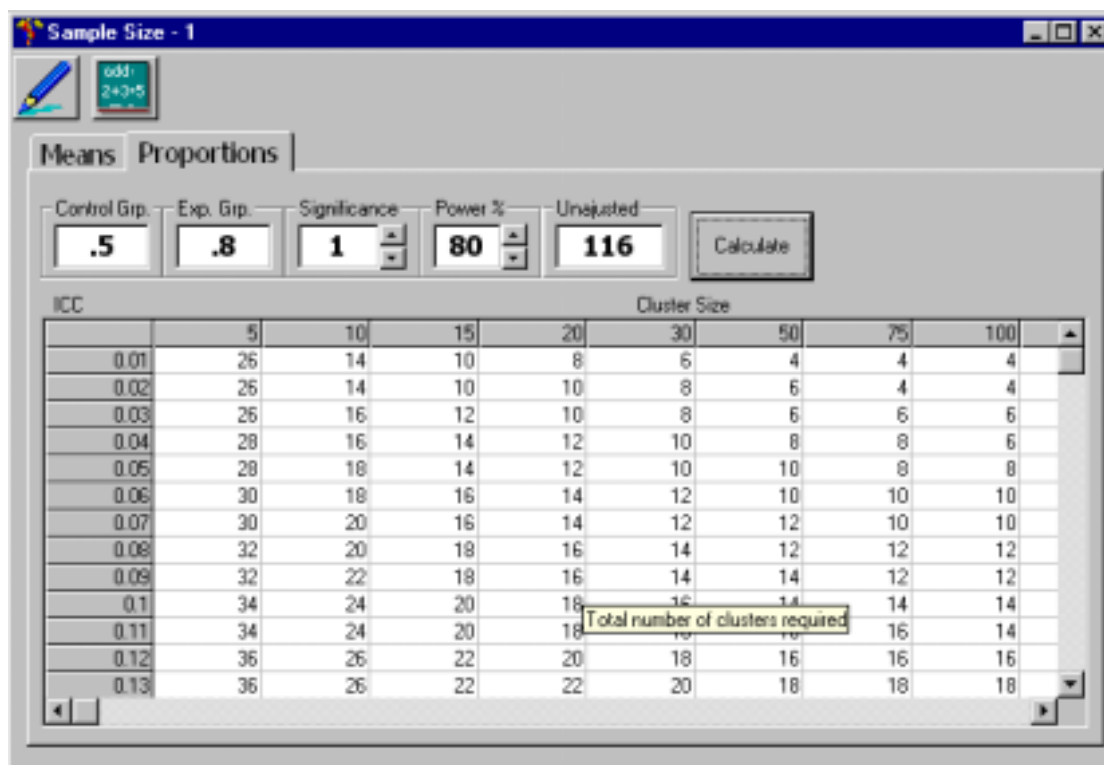
Proportions

For dichotomous data the outcome has only two categories, for example guideline compliance/non-compliance. Say we want to look at the difference in compliance rates for a guideline after the introduction of computer generated reminders. We have two groups, control and intervention. The intervention is randomised at the hospital level to minimise the risk of contamination from one group to the other. At baseline the compliance rate is 50%, and we are looking to see an improvement of 30% in the intervention group. The average cluster size is 23 and we have an estimation, say from literature, of 0.3 for the ICC in this case.

Click on *File* and select *New Sample Size*. When the window opens click on the *Proportions* tab. Enter:
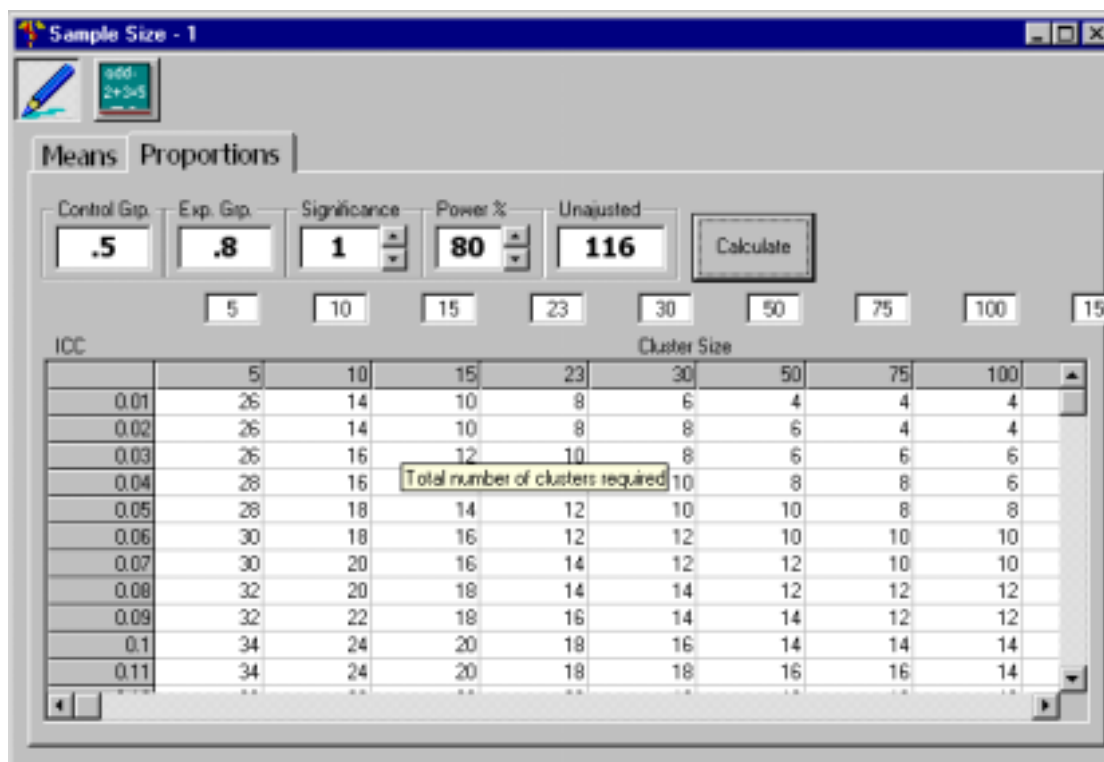- in the Control Grp. box (control group proportion) and
- in the Exp. Grp. box (experimental group proportion)

In this example we want a significance level of 0.01. Clicking on the down arrow to the right hand side of the Significance box will cycle the significance level through standard values. Alternatively we can click in the box and enter the new value.

Click on *Calculate*.

Sample Size - 1

Means | Proportions

Control Grp. **.5** | Exp. Grp. **.8** | Significance **1** | Power % **80** | Unajusted **116** | Calculate

ICC — Cluster Size

ICC	5	10	15	20	30	50	75	100
0.01	26	14	10	8	6	4	4	4
0.02	26	14	10	10	8	6	4	4
0.03	26	16	12	10	8	6	6	6
0.04	28	16	14	12	10	8	8	6
0.05	28	18	14	12	10	10	8	8
0.06	30	18	16	14	12	10	10	10
0.07	30	20	16	14	12	12	10	10
0.08	32	20	18	16	14	12	12	12
0.09	32	22	18	16	14	14	12	12
0.1	34	24	20	18	16	14	14	14
0.11	34	24	20	18			16	14
0.12	36	26	22	20	18	16	16	16
0.13	36	26	22	22	20	18	18	18

Total number of clusters required

To enter our average cluster size, 23, click on the *Pencil* icon to bring up the boxes for manual entry of cluster size. Select a box and type in 23. (It does not matter which box you select). Clicking on *calculate* again will adjust the values in that column for the change in cluster size.

Sample Size - 1

Means | Proportions

Control Grp. **.5** | Exp. Grp. **.8** | Significance **1** | Power % **80** | Unajusted **116** | Calculate

5 | 10 | 15 | 23 | 30 | 50 | 75 | 100 | 15

ICC — Cluster Size

ICC	5	10	15	23	30	50	75	100
0.01	26	14	10	8	6	4	4	4
0.02	26	14	10	8	8	6	4	4
0.03	26	16	12	10	8	6	6	6
0.04	28	16		10	10	8	8	6
0.05	28	18	14	12	10	10	8	8
0.06	30	18	16	12	12	10	10	10
0.07	30	20	16	14	12	12	10	10
0.08	32	20	18	14	14	12	12	12
0.09	32	22	18	16	14	14	12	12
0.1	34	24	20	18	16	14	14	14
0.11	34	24	20	18	18	16	16	14

Total number of clusters required

To find the appropriate number of clusters you may need to scroll down through the table to find the ICC of 0.3. The number of clusters needed is 40, an increase in sample size from 116 to 920, highlighting the significant effect of clustering.

Detectable Differences

The sample size calculator also calculates detectable differences for continuous and binary data. Click on *file* and select *New Difference Detectable.*

Means

For continuous data select *Means.* The Cluster Size , Number of Clusters and ICC are user defined. The sample size calculator will calculate the proportion of the standard deviation detectable, (the standardised difference δ/σ) for differing levels of significance and power, should a cluster randomised trial be run under these conditions.

For example, a cluster randomised trial is being planned with 10 clusters of average size 25 (with an assumed ICC of 0.01). We want to know the minimum clinical difference that such a trial can detect. First of all click on *File* and select *New Difference Detectable.* Enter the cluster size and number of clusters in the appropriate boxes. Click on *calculate.*

Along the top of the table are the power levels and down the side are the significance levels. The cells in the table contain the proportion of standard deviation detectable for specific levels of significance and power. From the table we can see that for 5% significance and 80% power, the standardised difference is 0.394 (i.e. 0.394 of a standard deviation can be detected). To express this in natural measurement units, multiply this value by the estimated standard deviation.

Proportions

For binary data select the *Proportions* tab on the *Detectable Differences* window. In the binary case there is a box to enter the expected control group proportion. The cells in the table contain the proportion detectable in the experimental group, when we are looking for an *increase* from the control group proportion. In this situation we are looking at a positive outcome eg compliance rate.

The Cluster Size, Number of Clusters and an estimate of the ICC are entered as before and also the expected control group proportion. Clicking on *Calculate* will give the proportion expected in the experimental group and the difference between the two is the minimum increase that can be detected under these conditions.

(**Note**: if the calculator returns a value of 1 or above this indicates that the sample size is too small to detect meaningful differences)

Because the sample size calculator is framed towards positive outcomes (where an increase in proportion is desirable), a slightly different approach is required if we are looking to detect a decrease of some undesirable outcome eg mortality.

For example, a clinical guideline is formulated to minimise inappropriate prescribing for sore throat. A cluster randomised trial at practice level is to be set up to examine the effectiveness of the guideline. Twelve practices are available with an average size of 20 patients. Current estimates of inappropriate or unnecessary prescribing are 70% (an ICC of 0.05 is assumed). Entering these figures into the sample size calculator will give an experimental proportion of approximately 0.90 (5% significance, 80% power). This is not framed in terms of a decrease. To get the associated decrease we must enter in the *Control Grp.* box $1 - p$, where p is the expected proportion in the control group, in this case 1-0.7 = 0.3. The value returned by the sample size calculator on this occasion for the detectable experimental group proportion is 0.54. This value needs to be reversed to express this in the terms of the original proportion (1-0.54=0.46). Hence the decrease detectable in inappropriate prescribing is from 0.7 down to 0.46.

References:

1. Kish L. *Survey sampling.* Wiley, NewYork, 1965.

2. Donner A, Birkett N, Buck C. Randomization by Cluster, Sample Size Requirements and Analysis, *American Journal of Epidemiology* 1981; **114**: 906-915.

To detect a difference δ, SSC uses the following equation to calculate the total unadjusted sample size:

$$m = \frac{4\left(z_{1-\alpha/2} + z_{1-\beta}\right)^2}{d^2}$$

where $d = \delta/\sigma$, the standardised difference, and σ the standard deviation. The sample size is then adjusted using the Variance Inflation Factor (or Design Effect) $1 + \left(\bar{n} - 1\right)\rho$, where $\bar{n}$ is the average cluster size. As we can see this factor is inflated by an increasing ICC and increasing average cluster size.

If we let p_A and p_B be the expected proportions in groups A and B, the formula SSC uses to calculate the unadjusted sample size for a binary outcome is

$$m = \frac{\left[z_{1-\alpha/2}\sqrt{2q(1-q)} + z_{1-\beta}\sqrt{p_A(1+p_A) + p_B(1+p_B)}\right]^2}{\left(p_A - p_B\right)^2},$$

where $q = \left(p_A + p_B\right)/2$. The VIF is then used as before to adjust the sample size for clustering.

For the detectable differences the SSC reverses the process. The sample size that it is input is adjusted for clustering and fed into the above formula rearranged to give the effect size. The binary formula leads to a quadratic that gives two roots. By default, SCC chooses the root that gives an increase from control to proportion, which is why a query about a decease must be rephrased slightly.