SAMPLE SIZE AND POWER

Formulation of the Problem

The problem is to determine the number of subjects required in order to be 100 x (1-ß)% of identifying an effect size equal to R or greater at the 100 x (1-a) % level of statistical significance.

The effect size R, depends on the particular study design and epidemiologic issue. It could be a difference between two means, a difference between two proportions, and odds ratio, a relative risk, etc.

The entity, 100 x (1-ß)%, represents the degree of certainty that the study will identify the risk factor, if it is as large or larger then R (Generally, 100 x (1-ß)% is set at 80% or 90%)

The entity, 100 x (1-a)%, represents the desired level of significance (Generally, 100 x(1-a)% set at 95%)

SAMPLE SIZE ESTIMATION FOR CASE CONTROL STUDIES

- **The sample size issue for unmatched case control studies having dichotomous exposures, dichotomous disease status, and no stratification is as follows: One needs to determine the number of cases and controls necessary to have 100 ´ (1 - b)% confidence of identifying a Relative Risk (exposure odds ratio) of R or larger at the 100 ´ (1 - a)% level of statistical significance if the rate of exposure in the controls is equal to p_0.**

- **It can be shown that the exposure rate, p_1 among cases would then be equal to:**

$$p_1 = \frac{p_0 R}{\left[1 + p_0\left(R - 1\right)\right]}$$

- **With this in mind, we have the following equation for the necessary cases and controls (corresponds to PEPI results without continuity correction factor)**

$$n = \frac{\left(z_a\sqrt{2\,\overline{p}\overline{q}} + z_b\sqrt{p_1 q_1 + p_0 q_0}\right)^2}{\left(p_1 - p_o\right)^2}$$

where

$$\overline{p} = \frac{p_1 + p_0}{2}$$

$$\overline{q} = 1 - \overline{p}$$

$$q_1 = 1 - p_1$$

$$q_0 = 1 - p_0$$

For 95% Significance, two tailed use z_a=1.96
For 80% Power, use $z_ß$=.84

Example of Sample Size Determination

Suppose you want to conduct a case control study to evaluate the nature of the putative association between exposure to cellular phones and incidence of acoustic neuroma (a benign but often troublesome tumor on the auditory nerve). The study should be designed in such a way that one would be 80% certain of estimating a relative risk (or odds ratio) that is at least twofold in magnitude, at a significance level of 95% (using a two tailed test).

The cases will be chosen from a registry of acoustic neuroma, and the controls will be selected by random digit dialing. The number of cases and controls will be the same.

It is projected that 10% of the population will be exposed to cellular phones.

FOR THIS EXAMPLE:

$$p_0 = .10$$

$$p_1 = \frac{.10 \times 2}{1 + (.10 \times 1)} = .1818$$

$$\overline{p} = \frac{.1818 + .10}{2} = .1409$$

$$\overline{q} = .8591$$

$$q_1 = .8182$$

$$q_0 = .90$$

$$n = \frac{\left[1.96\sqrt{2 \times .1409 \times .8591} + .84\sqrt{(.1818 \times .8182) + (.10 \times .9)}\right]^2}{(.1818 - .10)^2} =$$

$$: \frac{1.8902}{.0067} = 283$$

SAMPLE SIZE DETERMINATION USING PEPI-*SAMPLES*

```
SAMPLES - Sample Sizes for Comparison of Two Samples
ÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍ

      Significance level (two-tailed test)        = 0.05
      Power                                       = 80%
      Ratio of sample sizes (B:A)                 = 1
      Proportion in population B                  = 0.10
      Odds ratio of A to B                        = 2
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ

                          REQUIRED SAMPLE SIZE:
         Total 566: 283 in group A ("cases"), 283 in B ("controls")
                      or (with continuity correction):
         Total 614: 307 in group A ("cases"), 307 in B ("controls")

       EXPECTED PRECISION of observed difference between proportions:
       Approximate 95% C.I. = observed difference plus or minus 0.058
```

NOW CALCULATE SAMPLE SIZE FOR POWER OF 90%

```
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ

      Significance level (two-tailed test)        = 0.05
      Power                                       = 90%
      Ratio of sample sizes (B:A)                 = 1
      Proportion in population B                  = 0.10
      Odds ratio of A to B                        = 2
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ
                          REQUIRED SAMPLE SIZE:
         Total 756: 378 in group A ("cases"), 378 in B ("controls")
                      or (with continuity correction):
         Total 806: 403 in group A ("cases"), 403 in B ("controls")

       EXPECTED PRECISION of observed difference between proportions:
       Approximate 95% C.I. = observed difference plus or minus 0.050
```

NOW TRY POWER OF 70%

```
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ

      Significance level (two-tailed test)        = 0.05
      Power                                       = 70%
      Ratio of sample sizes (B:A)                 = 1
      Proportion in population B                  = 0.10
      Odds ratio of A to B                        = 2
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ

                          REQUIRED SAMPLE SIZE:
         Total 446: 223 in group A ("cases"), 223 in B ("controls")
                      or (with continuity correction):
         Total 494: 247 in group A ("cases"), 247 in B ("controls")

       EXPECTED PRECISION of observed difference between proportions:
       Approximate 95% C.I. = observed difference plus or minus 0.065
```

NOW TRY ORIGINAL DATA, BUT USE OR of 3

ÄÄÄ

```
      Significance level (two-tailed test)        = 0.05
      Power                                       = 80%
      Ratio of sample sizes (B:A)                 = 1
      Proportion in population B                  = 0.10
      Odds ratio of A to B                        = 3
```

ÄÄÄ

REQUIRED SAMPLE SIZE:
Total 200: 100 in group A ("cases"), 100 in B ("controls")
or (with continuity correction):
Total 226: 113 in group A ("cases"), 113 in B ("controls")

EXPECTED PRECISION of observed difference between proportions:
Approximate 95% C.I. = observed difference plus or minus 0.106

NOW TRY OR OF 1.5

ÄÄÄ

```
      Significance level (two-tailed test)        = 0.05
      Power                                       = 80%
      Ratio of sample sizes (B:A)                 = 1

      Proportion in population B                  = 0.10
      Odds ratio of A to B                        = 1.5
```

ÄÄÄ

REQUIRED SAMPLE SIZE:
Total 1,822: 911 in group A ("cases"), 911 in B ("controls")
or (with continuity correction):
Total 1,914: 957 in group A ("cases"), 957 in B ("controls")

EXPECTED PRECISION of observed difference between proportions:
Approximate 95% C.I. = observed difference plus or minus 0.030

NOW TRY OR OF 1.5 BUT WITH TWO CONTROLS PER CASE

ÄÄÄ

```
      Significance level (two-tailed test)        = 0.05
      Power                                       = 80%
      Ratio of sample sizes (B:A)                 = 2
      Proportion in population B                  = 0.10
      Odds ratio of A to B                        = 1.5
```

ÄÄÄ

REQUIRED SAMPLE SIZE:
Total 2,007: 669 in group A ("cases"), 1,338 in B ("controls")
or (with continuity correction):
Total 2,112: 704 in group A ("cases"), 1,408 in B ("controls")

EXPECTED PRECISION of observed difference between proportions:
Approximate 95% C.I. = observed difference plus or minus 0.030

Formulas for Multiple Controls per Case

c=# of controls per case

$$n = \frac{\left[z_a \sqrt{\left(1+\frac{1}{c}\right)\overline{p}'\overline{q}'} + z_b \sqrt{p_1 q_1 + \frac{p_0 q_0}{c}} \right]^2}{\left(p_1 - p_0\right)^2}$$

where

$$\overline{p}' = \frac{\left(p_1 + c p_0\right)}{1+c}$$

$$\overline{q}' = 1 - \overline{p}'$$

$$p_1 = \frac{p_0 R}{\left[1 + p_0\left(R-1\right)\right]}$$

n=# of cases needed
c x n=# of controls needed

EFFECT OF CONFOUNDING ON REQUIRED SAMPLE SIZES

Presence of confounding variables will require a larger sample size. Since the confounding variables must be controlled for in the analysis, a more complex statistical model must be fit, thus requiring a larger sample to achieve the stated specifications.

CALCULATING STUDY POWER

If financial constraints limit the study size (i.e. the number of cases and controls is fixed), can determine study power using varying values of R or z_a

Change values in SAMPLES—i.e. change significance level, ratio of controls to cases, Odds Ratio you want to detect, etc

Can also use POWER to determine if in a study already completed, what the power of your study was to detect different levels of R at a fixed significance level.

If you found a non-significant elevated OR or RR, you can determine if your study had adequate POWER to detect such a risk. Failure to reject the Null could be due to low power (i.e. inadequate sample size)

FORMULA FOR CALCULATING POWER—
Equal number of cases, controls
From Schlesselman

$$z_b = \frac{\left[\sqrt{n\left(p_1 - p_0\right)^2} - z_a\sqrt{2\overline{p}\overline{q}}\right]}{\sqrt{p_1 q_1 + p_0 q_0}}$$

$p_0 =$ **proportion exposed in controls (use population estimate)**

$$p_1 = \frac{p_0 R}{1 + p_0 (R - 1)}$$

$$\overline{p} = \frac{p_1 + p_0}{2}$$

$$\overline{q} = 1 - \overline{p}$$

$$q_1 = 1 - p_1$$

$$q_0 = 1 - p_0$$

n=number of cases, controls in each group

Find power corresponding to 1-$Z_ß$

EXAMPLE FROM SCHLESSELMAN, p.149:

For study with n=50, p_0=30 and a=.05 (two sided), what is power to detect a twofold increased risk?

See p. 149

Power= 38%

POWR MODULE IN PEPI

Use Option 1 for Independent samples (i.e. not matched)—this is for case control studies and cohort studies with exposed and non-exposed groups

Enter level of significance desired

Choose Option A to enter size of sample A and ratio of B to A
(For case control studies, A is number of cases, B is ratio of controls to cases)
(For cohort studies, A is size of exposed group, B is ratio of exposed to unexposed)

Enter size of sample A and ratio

Enter proportion in Sample B—
(For case control studies, proportion is proportion exposed in controls)
(For cohort studies, proportion is proportion of non-exposed with disease)

Then choose option A for cohort studies (rate ratio)
Option B for case control studies (odds ratio)
And enter effect that you want to detect (e.g. Odds ratio of 1.5)

USE OF *POWR* MODULE FOR STUDY DESIGN—OR FOR DETERMINING POWER OF COMPLETED STUDY

Say you wish to estimate an odds ratio of 1.5 with statistical power equal to 80% at significance level of 5%:

1. Try 250 cases with 1 control per case

```
POWR  -  Power of a Test Comparing Two Samples

ííííííííííííííííííííííííííííííííííííííííííííííííííííííííííííííííííííííííííííííí
      Significance level (two-tailed test)         = 0.05
      Sample size, population A                    = 250
      Ratio of sample B to sample A                = 1
      Proportion in population B                   = 0.15
      Odds ratio of A to B                         = 1.5
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ

                                POWER = 36.14%

ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ
```

2. Try 2 controls per case, 250 cases

```
      Significance level (two-tailed test)         = 0.05
      Sample size, population A                    = 250
      Ratio of sample B to sample A                = 2
      Proportion in population B                   = 0.15
      Odds ratio of A to B                         = 1.5
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ

                                POWER = 48.98%
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ
```

3. Try 3 controls per case, 250 cases

```
Significance level (two-tailed test)         = 0.05
      Sample size, population A                    = 250
      Ratio of sample B to sample A                = 3
      Proportion in population B                   = 0.15
      Odds ratio of A to B                         = 1.5
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ

                                POWER = 54.78%
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ
```

4. Try 4 controls per case, 250 cases

```
      Significance level (two-tailed test)         = 0.05
      Sample size, population A                    = 250
      Ratio of sample B to sample A                = 4
      Proportion in population B                   = 0.15
      Odds ratio of A to B                         = 1.5
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ

                                POWER = 58.04%
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ
```

5. Bite bullet and increase cases to 300

```
      Significance level (two-tailed test)         = 0.05
      Sample size, population A                    = 300
      Ratio of sample B to sample A                = 4
      Proportion in population B                   = 0.15
      Odds ratio of A to B                         = 1.5
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ

                              POWER = 66.00%

ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ
```

6. Still not good enough, increase cases to 325

```
      Significance level (two-tailed test)         = 0.05
      Sample size, population A                    = 325
      Ratio of sample B to sample A                = 4
      Proportion in population B                   = 0.15
      Odds ratio of A to B                         = 1.5
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ

                              POWER = 69.51%

ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ
```

7. Increase cases to 350

```
      Significance level (two-tailed test)         = 0.05
      Sample size, population A                    = 350
      Ratio of sample B to sample A                = 4
      Proportion in population B                   = 0.15
      Odds ratio of A to B                         = 1.5
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ

                              POWER = 72.71%
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ
```

8. Try 400 Cases

```
      Significance level (two-tailed test)         = 0.05
      Sample size, population A                    = 400
      Ratio of sample B to sample A                = 4
      Proportion in population B                   = 0.15
      Odds ratio of A to B                         = 1.5
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ
                              POWER = 78.28%
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ
```

9. Almost, let's see if 425 will do it

```
     Significance level (two-tailed test)         = 0.05
     Sample size, population A                    = 425
     Ratio of sample B to sample A                = 4
     Proportion in population B                   = 0.15
     Odds ratio of A to B                         = 1.5
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ

                          POWER = 80.68%
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ
```

EFFECT OF VARYING SPECIFICATIONS

1. All else being equal, it will take a larger (smaller) sample to detect a smaller (larger) effect size.

Number, n, of cases and controls in each group required for various values of Odds Ratio, R, for a=.10, ß=.20, and p_0=.10 (for equal number of cases, controls)

Odds Ratio, R	Number, n, of cases and controls
1.20	3850
1.50	718
1.70	401
2.00	223
2.50	119
3.00	79
4.00	46
5.00	32
10.00	14

2. All else being equal, the greater the background exposure prevalence or prevalence of non-exposure, the smaller the sample size required

Odds ratio=2, a=.10, ß=.20

Background prevalence	Number, n, of cases and controls
.01	1889
.05	406
.10	223
.20	135
.30	111
.40	105
.50	108
.60	120
.70	147

Effect of Other Parameters on Required Sample Size

Raising the Confidence Level Parameter 100 x (1-ß)% will increase the sample size required (the more power required, the larger the sample size required)

Lowering the significance level, a, (e.g., from .05 to .01) will cause an increase in the necessary sample size.

If number of cases is limited (i.e. not a common disease) can increase number of controls per case to increase power. Incremental gain after 4 controls per case however is limited.

EFFECT OF NUMBER OF CONTROLS PER CASE ON THE RELIABILITY OF THE RESULTING ESTIMATES

Controls Per Case	Reliability of Resulting Odds Ratio (Relative to One Control Per Case	Incremental Gain
1	1.00	-
2	1.33	33%
3	1.50	17%
4	1.60	10%
5	1.67	7%
6	1.71	4%
7	1.76	3%

VERY SMALL INCREMENTAL GAIN AFTER 4 CONTROLS PER CASE

EFFECT OF NUMBER OF CONTROLS PER CASE ON POWER—USING a=.10, ß=.20, Odds Ratio=2

Controls Per Case	POWER
1	76.2
2	86.5
3	89.9
4	91.5
5	92.4
6	93.0
7	93.4

SAMPLE SIZE FOR MATCHED STUDIES

To determine number of total pairs required, need to know probability of an exposure-discordant pair.

Can estimate using estimated proportion of exposed controls in target population, and estimated proportion of exposed cases.

Need to estimate similarity of findings based on matching –i.e. what is the extent to which matching will make the findings in the two samples similar, necessitating larger sample sizes (i.e. if there are a lot of concordant pairs, you will need a larger number of matched pairs in order to get a sufficient number of discordant pairs). The matching factor is seldom more than 2.5.

Can use SAMPLES in PEPI—USE OPTION 2 for MATCHED STUDIES

```
SAMPLES - Sample Sizes for Comparison of Two Samples
Tuesday, 8th October 2002.
ÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍ

      Significance level (two-tailed test)         = 0.05
      Power                                        = 80%
      Proportion in population B                   = 0.10
      Odds ratio (based on discordant pairs)       = 2
      Effect-of-matching factor                    = 2.5
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ

                            REQUIRED SAMPLE SIZE:

                 340 pairs (estimated discordant pairs: 71)
      or 255 case-and-2-controls sets; or 227 case-and-3-controls sets
      or 213 case-and-4-controls sets; or 204 case-and-5-controls sets

ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ


SAMPLES - Sample Sizes for Comparison of Two Samples
Tuesday, 8th October 2002.
ÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍ

      Significance level (two-tailed test)         = 0.05
      Power                                        = 80%
      Proportion in population B                   = 0.10
      Odds ratio (based on discordant pairs)       = 2
      Effect-of-matching factor                    = 1
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ

                            REQUIRED SAMPLE SIZE:

                 286 pairs (estimated discordant pairs: 71)
      or 215 case-and-2-controls sets; or 191 case-and-3-controls sets
      or 179 case-and-4-controls sets; or 172 case-and-5-controls sets

ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ
```

SAMPLE SIZE FOR COHORT STUDIES

Difference in disease proportions between two groups

See handout for formulas

FOR PEPI SAMPLES:

Need proportion with disease in non-exposed, proportion exposed, and ratio of exposed group to non-exposed

USE option 1

Enter Significance level, power, proportion of disease in unexposed group, ratio of exposed to unexposed

Enter 1 when prompted for difference to be detected (rate ratio)

Enter Relative Risk to be detected

SAMPLE SIZE FOR COHORT STUDY
A=Exposed, B=Unexposed

ASSUME TWICE AS MANY UNEXPOSED AS EXPOSED

```
SAMPLES - Sample Sizes for Comparison of Two Samples
Monday, 7th October 2002.
ÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍÍ
      Significance level (two-tailed test)        = 0.05
      Power                                       = 80%
      Ratio of sample sizes (B:A)                 = 2
      Proportion in population B                  = 0.10
      Rate ratio of A to B                        = 2
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ

                            REQUIRED SAMPLE SIZE:
         Total 432: 144 in group A ("cases"), 288 in B ("controls")
                      or (with continuity correction):
         Total 474: 158 in group A ("cases"), 316 in B ("controls")

       EXPECTED PRECISION of observed difference between proportions:
       Approximate 95% C.I. = observed difference plus or minus 0.069
```

NOW USE SAME NUMBER OF UNEXPOSED AS EXPOSED

```
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ

      Significance level (two-tailed test)        = 0.05
      Power                                       = 80%
      Ratio of sample sizes (B:A)                 = 1
      Proportion in population B                  = 0.10
      Rate ratio of A to B                        = 2
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ

                            REQUIRED SAMPLE SIZE:
         Total 398: 199 in group A ("cases"), 199 in B ("controls")
                      or (with continuity correction):
         Total 438: 219 in group A ("cases"), 219 in B ("controls")

       EXPECTED PRECISION of observed difference between proportions:
       Approximate 95% C.I. = observed difference plus or minus 0.071
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ


      Significance level (two-tailed test)        = 0.05
      Power                                       = 80%
      Ratio of sample sizes (B:A)                 = 1
      Proportion in population B                  = 0.10
      Rate ratio of A to B                        = 2
      Ratio of sample sizes (B:A)                 = 1
      Proportion in population B                  = 0.1
      Rate ratio of A to B                        = 2
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ
```

IN A COHORT STUDY, FOR FIXED NUMBER OF EXPOSED, CAN INCREASE POWER BY INCREASING RATIO OF UNEXPOSED TO EXPOSED

```
POWR  -  Power of a Test Comparing Two Samples
Tuesday, 8th October 2002.
-----------------------------------------------------------------------

      Significance level (two-tailed test)         = 0.05
      Sample size, population A                    = 200
      Ratio of sample B to sample A                = 1
      Proportion in population B                   = 0.10
      Rate ratio of A to B                         = 2
-----------------------------------------------------------------------

                  POWER = 80.20% (no continuity correction)
                       or 75.92% (with continuity correction)

-----------------------------------------------------------------------

      Significance level (two-tailed test)         = 0.05
      Sample size, population A                    = 200
      Ratio of sample B to sample A                = 2
      Proportion in population B                   = 0.10
      Rate ratio of A to B                         = 2
-----------------------------------------------------------------------


                  POWER = 90.68% (no continuity correction)
                       or 88.53% (with continuity correction)

-----------------------------------------------------------------------

      Significance level (two-tailed test)         = 0.05
      Sample size, population A                    = 200
      Ratio of sample B to sample A                = 3
      Proportion in population B                   = 0.10
      Rate ratio of A to B                         = 2
-----------------------------------------------------------------------


                  POWER = 93.67% (no continuity correction)
                       or 92.18% (with continuity correction)

-----------------------------------------------------------------------

      Significance level (two-tailed test)         = 0.05
      Sample size, population A                    = 200
      Ratio of sample B to sample A                = 4
      Proportion in population B                   = 0.10
      Rate ratio of A to B                         = 2
-----------------------------------------------------------------------

                  POWER = 94.99% (no continuity correction)
                       or 93.80% (with continuity correction)
```

```
--------------------------------------------------------------------

      Significance level (two-tailed test)         = 0.05
      Sample size, population A                    = 200
      Ratio of sample B to sample A                = 5
      Proportion in population B                   = 0.10
      Rate ratio of A to B                         = 2
-------------------------------------------------------------------


                  POWER = 95.71% (no continuity correction)
                       or 94.70% (with continuity correction)

--------------------------------------------------------------------

      Significance level (two-tailed test)         = 0.05
      Sample size, population A                    = 200
      Ratio of sample B to sample A                = 6
      Proportion in population B                   = 0.10
      Rate ratio of A to B                         = 2
--------------------------------------------------------------------


                  POWER = 96.17% (no continuity correction)
                       or 95.26% (with continuity correction)

--------------------------------------------------------------------

      Significance level (two-tailed test)         = 0.05
      Sample size, population A                    = 200
      Ratio of sample B to sample A                = 7
      Proportion in population B                   = 0.10
      Rate ratio of A to B                         = 2
--------------------------------------------------------------------


                  POWER = 96.48% (no continuity correction)
                       or 95.64% (with continuity correction)

--------------------------------------------------------------------
```

FOR FIXED TOTAL NUMBER OF EXPOSED AND UNEXPOSED, NO GAIN IN POWER BY INCREASING NUMBER OF UNEXPOSED, REDUCING NUMBER OF EXPOSED

```
POWR  -  Power of a Test Comparing Two Samples
Tuesday, 8th October 2002.
----------------------------------------------------------------------

      Significance level (two-tailed test)         = 0.05
      Combined sample size                         = 500
      Ratio of sample B to sample A                = 1
      Proportion in population B                   = 0.10
      Rate ratio of A to B                         = 2
----------------------------------------------------------------------

                  POWER = 88.16% (no continuity correction)
                       or 85.40% (with continuity correction)

----------------------------------------------------------------------
      Significance level (two-tailed test)         = 0.05
      Combined sample size                         = 500
      Ratio of sample B to sample A                = 2
      Proportion in population B                   = 0.10
      Rate ratio of A to B                         = 2
----------------------------------------------------------------------

                  POWER = 85.29% (no continuity correction)
                       or 82.06% (with continuity correction)

----------------------------------------------------------------------

      Significance level (two-tailed test)         = 0.05
      Combined sample size                         = 500
      Ratio of sample B to sample A                = 3
      Proportion in population B                   = 0.10
      Rate ratio of A to B                         = 2
----------------------------------------------------------------------


                  POWER = 80.17% (no continuity correction)
                       or 76.03% (with continuity correction)

----------------------------------------------------------------------

      Significance level (two-tailed test)         = 0.05
      Combined sample size                         = 500
      Ratio of sample B to sample A                = 4
      Proportion in population B                   = 0.10
      Rate ratio of A to B                         = 2
----------------------------------------------------------------------

                  POWER = 74.98% (no continuity correction)
                       or 69.94% (with continuity correction)
```

```
------------------------------------------------------------------------

      Significance level (two-tailed test)         = 0.05
      Combined sample size                         = 500
      Ratio of sample B to sample A                = 5
      Proportion in population B                   = 0.10
      Rate ratio of A to B                         = 2
------------------------------------------------------------------------


                  POWER = 70.20% (no continuity correction)
                       or 64.36% (with continuity correction)

-----------------------------------------------------------------------

      Significance level (two-tailed test)         = 0.05
      Combined sample size                         = 500
      Ratio of sample B to sample A                = 6
      Proportion in population B                   = 0.10
      Rate ratio of A to B                         = 2
------------------------------------------------------------------------


                  POWER = 65.92% (no continuity correction)
                       or 59.40% (with continuity correction)

------------------------------------------------------------------------

      Significance level (two-tailed test)         = 0.05
      Combined sample size                         = 500
      Ratio of sample B to sample A                = 7
      Proportion in population B                   = 0.10
      Rate ratio of A to B                         = 2
-----------------------------------------------------------------------


                  POWER = 62.14% (no continuity correction)
                       or 55.03% (with continuity correction)

-----------------------------------------------------------------------
```