Chapter 276

Tests for Paired Sensitivities

Introduction

This procedure gives power or required sample size for comparing two diagnostic tests when the outcome is sensitivity. In the design covered in this chapter, the sensitivities of two diagnostic tests are each performed on the same subject.

Specifically, a set of *N* subjects is selected and each is given the both of the diagnostic tests being considered. Subsequently, a gold standard test is also given and used to obtain the true presence or absence of the disease. The gold standard may be a more expensive test or require the sacrifice of the subject.

Measures of diagnostic accuracy are sensitivity and specificity. *Sensitivity* (Se) is the probability that the diagnostic test is positive for the disease, given that the subject actually has the disease. *Specificity* (Sp) is the probability that the diagnostic test is negative, given that the subject does not have the disease. Mathematically,

Sensitivity = Pr(+test|disease)

Specificity = Pr(-test|no disease)

Li and Fine (2004) present sample size methodology for testing sensitivity and specificity using a paired, prospective design. Their methodology is used here. Other useful references are Obuchowski and Zhou (2002), Machin, Campbell, Tan, and Tan (2009), and Zhou, Obuchowski, and McClish (2002).

Technical Details

In this, two diagnostic tests are administered to each of N subjects. Of those, N1 have the disease of interest. The N2 subjects without the disease are set aside and only the N1 diseased subjects are used in the analysis. The results can be displayed in the following 2-by-2 table.

	Test 2		
Test 1	**Yes**	**No**	**Total**
Yes	a	b	a+b
No	c	d	c+d
Total	a+c	b+d	N1

McNemar's Test

The McNemar test statistic is

$$\chi_1^2 = \frac{(b-c)^2}{b+c}$$

The null hypothesis that the sensitivities of the two tests are equal, that P(a+b) = P(a+c) is equivalent to the hypothesis that P(b) = P(c). The two-sided alternative hypothesis is that P(b) ≠ P(c).

Power Analysis

The sample size problem reduces to a study of how many Yes-No's and No-Yes's are needed. Once this has been determined, the overall sample size is found using the proportion of discordant pairs to inflate the sample size appropriately.

Some power analysis programs follow an approximate procedure. Since the McNemar statistic follows the binomial probability distribution for a fixed number of discordant pairs, they use formulas that use the normal approximation to the binomial and then adjust the sample size depending on the proportion of discordant pairs, D=P(b)+P(c). This is called the conditional procedure.

One such approximate formula is given by Machin, Campbell, Fayers, and Pinol (1997).

$$N1 = \frac{\left(z_{1-\alpha/s}(\psi+1) + z_{1-\beta}\sqrt{(\psi+1)^2 - (\psi-1)^2 D} \right)^2}{(\psi-1)^2 D}$$

where s is the number of sides to the test (one or two), $\psi = \dfrac{P(b)}{P(c)}$, $D = P(b) + P(c)$, and α and β are the usual type 1 and type 2 error rate probabilities.

However, Schork and Williams (1980) published a formula which provides the exact results for the unconditional case. This is the formulation that is used in **PASS**.

$$Power = \sum_{x=r}^{N1}\sum_{y=0}^{I}\frac{N1!}{(N1-x)!\,y!\,(x-y)!}(1-D)^{N1-x}(D+\delta)^y(D-\delta)^{x-y}$$

where

$D = P(b) + P(c)$

$\delta = P(b) - P(c)$

N1 is total of all four cells

r is the smallest integer for which $\left(\dfrac{1}{2}\right)^r \le \alpha$

I is the largest integer such that $\sum_{j=0}^{IR}\binom{x}{j}\left(\dfrac{1}{2}\right)^x \le \alpha$

Finally, to obtain the estimate of N, we inflate N1 by the prevalence using N = N1/P, where P is the prevalence.

Specificity

This procedure computes the sample size for sensitivity. If you want a power analysis or sample size for specificity, you can use this procedure with the following minor adjustments.

1. Replace *Se* with *Sp* in all entries.

2. Replace the disease prevalence with 1 – prevalence. That is, the prevalence becomes the proportion without the disease.

Procedure Options

This section describes the options that are specific to this procedure. These are located on the Design tab. For more information about the options of other tabs, go to the Procedure Window chapter.

Design Tab

The Design tab contains the parameters associated with this test such as the sensitivities, specificities, sample sizes, alphas, and powers.

Solve For

Solve For

Select the parameter to be solved for in terms of the other parameters. Note that this parameter will be displayed on the vertical axis of the plot.

Possible choices for this parameter are

- **Se2 (Search > Se1)**

  Search for the Se2 (sensitivity of test 2 under H1) that results in the specified values of the other parameters. The binary search is conducted between Se1 and Min(1, Se1 + D).

- **Se2 (Search > Se1)**

  Search for the Se2 (sensitivity of test 2 under H1) that results in the specified values of the other parameters. The binary search is conducted between Max(0, Se1 - D) and Se1.

- **Power**

  Calculate the power based on the other parameters.

- **Sample Size**

  Search for the sample size needed to attain the values of the other parameters.

Test

Alternative Hypothesis (H1)

Specify the alternative hypothesis of the test. Usually, the two-sided option is selected.

Power and Alpha

Power

This option specifies one or more values for the desired power. Power is the probability of rejecting a false null hypothesis, and is equal to 1- Beta. Beta is the probability of a type-II error, which occurs when a false null hypothesis is not rejected.

Values must be between zero and one. Historically, the value of 0.80 (Beta = 0.20) was used for power. Now, 0.90 (Beta = 0.10) is commonly used.

A single value may be entered or a range of values such as *0.8 to 0.95 by 0.05* may be entered.

Alpha

This option specifies one or more values for the probability of a type-I error. A type-I error occurs when a true null hypothesis is rejected. For this procedure, a type-I error occurs when you reject the null hypothesis of equal sensitivities when in fact they are equal.

Values must be between zero and one. Historically, the value of 0.05 has been used for alpha and this is still the most common choice today.

Note that because of the discrete nature of the binomial distribution, the alpha level rarely will be achieved exactly.

A single value may be entered here or a range of values such as *0.05 to 0.2 by 0.05* may be entered.

Sample Size

N (Number of Subjects)

Enter a value (or range of values) for the sample size, N. N is the total number of subjects (both diseased and non-diseased) in the study.

You may enter a range of values such as 10 to 100 by 10.

P (Prevalence)

Specify one or more values for the disease prevalence: the anticipated proportion of the population of interest that has the disease. Because this is a proportion all values must be between zero and one.

You may enter a single value or a range of values such as *0.1, 0.2, 0.3*.

Effect Size

Se1 (Sensitivity of Test 1)

Enter the value of Se1, the sensitivity of both tests assumed by the null hypothesis, H0. The difference detected by this design is Se1 - Se2. The values must be between 0 and 1 and cannot be equal to Se2.

Sensitivity = Pr(+Test|Disease).

You may enter a range of values such as *0.7, 0.8, 0.9* or *0.6 to 0.9 by 0.1*.

Se2 (Sensitivity of Test 2)

Enter the value of Se2 which is the sensitivity in test 2 assumed by the alternative hypothesis, H1. The difference detected by this design is Se1 - Se2. The values must be between 0 and 1 and cannot be equal to Se1.

Note: Sensitivity = Pr(+Test|Disease).

You can enter a range of values such as *0.75, 0.85, 0.95* or *0.55 to 0.95 by 0.1*.

Proportion Discordant (P10+P01)

This is the proportion of discordant pairs D. This value will be difficult to specify unless you have previous studies that give you some idea of what to expect. When you have no idea, Machin, Campbell, Fayers, and Pinol (1997) suggest the following strategy. Estimate values of *Se1* and *Se1*. Calculate the proportion of discordant pairs using the approximation

$$D = Se1(1 - Se2) + Se2(1 - Se1)$$

This approximation assumes that the two responses are independent within each subject, which will usually not be true. However, it may be the only way of determining a ball park value for this parameter.

Options Tab

The Options tab allows the specification of options that control the calculations.

Approximations

Use Approximations if N(P) is greater than

Below this value of N(P), the exact power calculation formula based on the binomial is used. Above this value of N(P), the approximate formula based on the normal approximation to the binomial is used. The exact formula suffers from numerical problems when N(P) is greater than 2000. On the other hand, the approximate formula tends to underestimate the N(P) necessary to achieve a certain beta value by about 5%.

You control which formula is used by setting this value.

Example 1 – Finding the Power

Suppose that diagnosing a certain disease has used a certain diagnostic test which has a sensitivity of 71%. A new diagnostic test has been developed that is much less expensive and invasive. Researchers want to design a paired, prospective study in which each subject will be given both the old and new diagnostic tests. The data will be analyzed using a two-sided McNemar's test with a significance level of 0.05. From previous studies, the percent discordant is estimated at 30%.

They want to consider changes in sensitivity of 10%, 15%, 20%, and 25%. These changes translate to sensitivities of 79.20%, 81.65%, 85.20%, and 88.75%. The prevalence of the disease in the population of interest is 20%. The power will be determined for trials with sample sizes between 300 and 2400 incremented by 300.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Tests for Paired Sensitivities** procedure window by expanding **Proportions**, clicking on **Sensitivity and Specificity**, and then clicking on **Tests for Paired Sensitivities**. You may then make the appropriate entries as listed below, or open **Example 1** by going to the **File** menu and choosing **Open Example Template**.

Option	Value
Design Tab	
Solve For	**Power**
Alternative Hypothesis	**Two-Sided**
Alpha..	**0.05**
N (Number of Subjects)	**300 to 2400 by 300**
P (Prevalence)	**0.2**

Tests for Paired Sensitivities

Design Tab (continued)
Se1 (Sensitivity of Test 1)...................... **0.71**
Se2 (Sensitivity of Test 2)...................... **0.792 0.8165 0.852 0.8875**
D (Proportion Discordant) **0.3**

Annotated Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Results

Numeric Results for a Two-Sided, McNemar's Test

Power	N	Sensitivity of Test 1 Se1	Sensitivity of Test 2 Se2	Difference Se1-Se2	Proportion Discordant D	Prev- alence- P	Alpha	Beta
0.15137	300	0.710	0.792	-0.082	0.300	0.200	0.05000	0.84863
0.31319	600	0.710	0.792	-0.082	0.300	0.200	0.05000	0.68681
0.47378	900	0.710	0.792	-0.082	0.300	0.200	0.05000	0.52622
0.60339	1200	0.710	0.792	-0.082	0.300	0.200	0.05000	0.39661
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.

References
Obuchowski, N.A., Zhou, X.H. 2002. 'Prospective studies of diagnostic test accuracy when disease prevalence is low,' Biostatistics, Volume 3, No. 4, pages 477-492.
Li, J., Fine, J. 2004. 'On sample size for sensitivity and specificity in prospective diagnostic accuracy studies,' Statistics in Medicine, Volume 23, pages 2537-2550.
Machin, D., Campbell, M.J., Tan, S.B., Tan, S.H. 2009. Sample Size Tables for Clinical Studies, Third Edition. Wiley-Blackwell, Chichester, United Kingdom.
Zhou, X.H., Obuchowski, N.A., McClish, D.K. 2002. Statistical Methods in Diagnostic Medicine. Wiley-Interscience, New York.
Schork, M. and Williams, G. 1980. 'Number of Observations Required for the Comparison of Two Correlated Proportions.' Communications in Statistics-Simula. Computa., B9(4), 349-357.

Report Definitions
Power is the probability of rejecting a false null hypothesis. It should be close to one.
N is the total number of subjects in the study.
Se1 is the sensitivity of test 1 assuming both H0 and H1.
Se2 is the sensitivity for test 2 assuming H1.
Difference is the Se1-Se2 assuming H1.
D, the proportion discordant, is proportion of diseased subjects that have unmatching test results.
Prevalence is proportion of diseased individuals in the population.
Alpha is the probability of rejecting a true null hypothesis, H0. It should be small.
Beta is the probability of accepting a false null hypothesis, H0. It should be small.

Summary Statements
A sample size of 300 subjects achieves 15% power to detect a difference of -0.082 between two diagnostic tests whose sensitivities are 0.710 and 0.792. This procedure uses a two-sided McNemar test with a significance level of 0.05000. The prevalence of disease in the population is 0.200. The proportion of discordant pairs is 0.300.

This report shows the values of each of the parameters, one scenario per row.

Power

This is the power of the test.

N

This is the total sample size. The number of diseased subjects is N(P).

Se1

This is the sensitivity of tests assuming H0. It is the sensitivity of test 1 under H1. The sensitivity is the proportion of diseased subjects that yield a positive test result.

Se2

This is the sensitivity of test 2 assuming H1.

Difference

This is the difference between the two sensitivities under H1. It is calculated as Se1 – Se2.

Proportion Discordant

D, the proportion discordant, is proportion of diseased subjects that have non-matching test results.

Prevalence

P, the prevalence, is proportion of the population that actually has the condition of interest (disease).
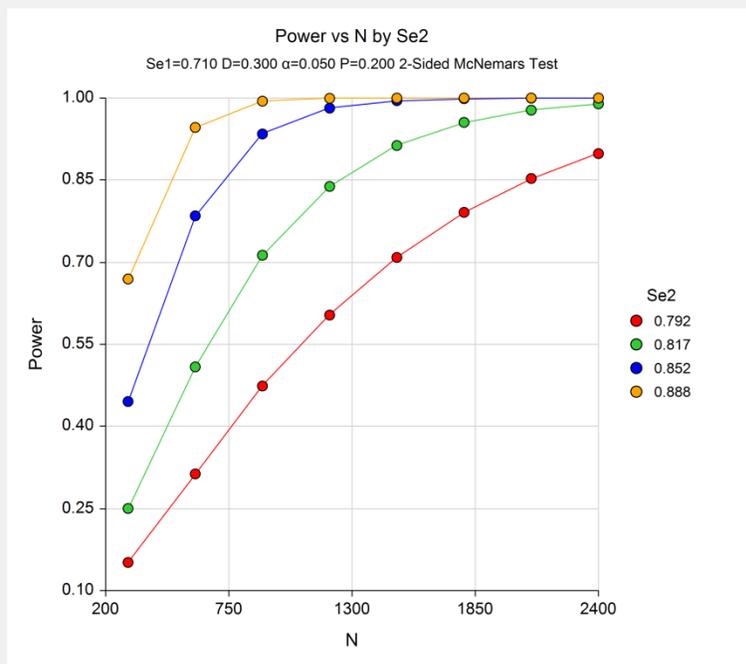
Alpha

This is the alpha (probability of rejecting H0 when H0 is true) that was desired.

Beta

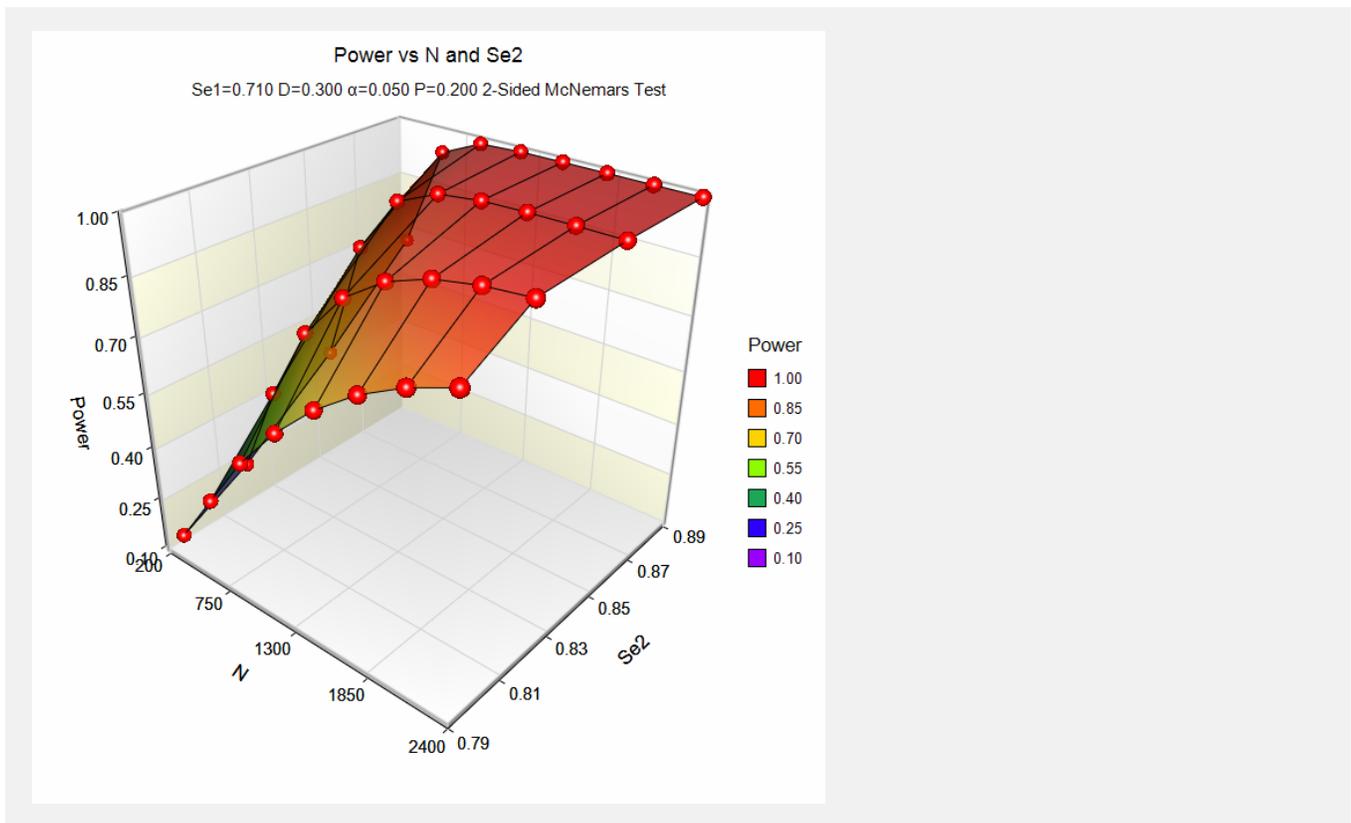Beta is the probability of accepting the null hypothesis when it is false.

Plots Section

Power vs N and Se2

Se1=0.710 D=0.300 α=0.050 P=0.200 2-Sided McNemars Test

These plots show the relationship between power, sample size, and Se2 in this example.

Example 2 – Finding the Sample Size

Continuing with Example 1, suppose you want to study the impact of various choices for Se2 on sample size. Using a significance level of 0.05 and 90% power, find the sample size when Se2 is 79.20%, 81.65%, 85.20%, and 88.75%. Assume a two-tailed test is used.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Tests for Paired Sensitivities** procedure window by expanding **Proportions**, clicking on **Sensitivity and Specificity**, and then clicking on **Tests for Paired Sensitivities**. You may then make the appropriate entries as listed below, or open **Example 2** by going to the **File** menu and choosing **Open Example Template**.

Option	Value
Design Tab	
Solve For	**Sample Size**
Alternative Hypothesis	**Two-Sided**
Power	**0.90**
Alpha	**0.05**
P (Prevalence)	**0.2**
Se1 (Sensitivity of Test 1)	**0.71**
Se2 (Sensitivity of Test 2)	**0.792 0.8165 0.852 0.8875**
D (Proportion Discordant)	**0.3**

Annotated Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Results

Numeric Results for a Two-Sided, McNemar's Test

Power	N	Sensitivity of Test 1 Se1	Sensitivity of Test 2 Se2	Difference Se1-Se2	Proportion Discordant D	Prev-alence- P	Alpha	Beta
0.90030	2415	0.710	0.792	-0.082	0.300	0.200	0.05000	0.09970
0.90097	1440	0.710	0.817	-0.107	0.300	0.200	0.05000	0.09903
0.90001	795	0.710	0.852	-0.142	0.300	0.200	0.05000	0.09999
0.90102	510	0.710	0.888	-0.178	0.300	0.200	0.05000	0.09898

It is interesting to note the reduction in sample size of the paired design over the two-sample design discussed in Chapter 275. A comparison of these sample sizes to those shows a reduction of over 50%.

Example 3 – Validation using Machin and Campbell (2009)

Machin, et al (2009) page 166-167 give the results of a sample size determination in which Se1 = 0.27, Se2 = 0.66, P = 0.25, alpha = 0.05 (two-sided), and power = 0.80. The resulting sample size is 80 using one formula and 100 using another. The formulas used in their book are based on a conditional argument that does not require the proportion discordant. We will compare our results to theirs using D = 0.4, 0.5, and 0.6.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Tests for Paired Sensitivities** procedure window by expanding **Proportions**, clicking on **Sensitivity and Specificity**, and then clicking on **Tests for Paired Sensitivities**. You may then make the appropriate entries as listed below, or open **Example 3** by going to the **File** menu and choosing **Open Example Template**.

Option	Value
Design Tab	
Solve For	**Sample Size**
Alternative Hypothesis	**Two-Sided**
Power..	**0.80**
Alpha...	**0.05**
P (Prevalence)	**0.25**
Se1 (Sensitivity of Test 1)......................	**0.66**
Se2 (Sensitivity of Test 2)......................	**0.27**
D (Proportion Discordant)	**0.4 0.5 0.6**

Annotated Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Results

Numeric Results for a Two-Sided, McNemar's Test

Power	N	Sensitivity of Test 1 Se1	Sensitivity of Test 2 Se2	Difference Se1-Se2	Proportion Discordant D	Prev-alence- P	Alpha	Beta
0.83196	80	0.660	0.270	0.390	0.400	0.250	0.05000	0.16804
0.80961	104	0.660	0.270	0.390	0.500	0.250	0.05000	0.19039
0.81101	128	0.660	0.270	0.390	0.600	0.250	0.05000	0.18899

PASS has also obtained an N of 80 for the case when D = 0.4. However, the sample size increases if D is 0.5 or 0.6.