Adjustment to the McNemar's Test for the Analysis of Clustered Matched-Pair Data

William F. McCarthy*

*Maryland Medical Research Institute, dr.w.f.mccarthy@gmail.com

Adjustment to the McNemar's Test for the Analysis of Clustered Matched-Pair Data

William F. McCarthy

Abstract

This paper presents how one can adjust the McNemar's test for the analysis of clustered matched-pair data. A McNemar's-like table for K clusters of matched-pair data is used.

The following discussion is based on the statistical procedure developed by Nancy A. Obuchowski of the Cleveland Clinic [1] and Valerie Durkalski [2] of the Medical University of South Carolina.

Obuchowski's approach avoids having to assume a constant within-cluster correlation (Durkalski's approach [2] does assume a constant within-cluster correlation).

Obuchowski's test statistic for clustered matched-pair data takes into account the possible variation in correlation between units within a cluster.

Obuchowski's approach considers the homogeneity hypothesis

H_0: $p_1 = p_2$ versus H_A: the p_i's are unequal

This is the same homogeneity hypothesis as used in the unadjusted McNemar's Test outlined by Conover [3].

Durkalski's approach considers the null hypothesis

$H_0 : p_{10k} - p_{01k} = 0$ for all K clusters versus $H_A : p_{10k} - p_{01k} \neq 0$

Table 1. McNemar's-Like Table for K Clusters of Matched-Pair Data.

Procedure 2	Procedure 1		
	Success	Failure	
Success	$\sum_{k=1}^{K} a_k$	$\sum_{k=1}^{K} b_k$	$\sum_{k=1}^{K} (a_k + b_k)$
Failure	$\sum_{k=1}^{K} c_k$	$\sum_{k=1}^{K} d_k$	
	$\sum_{k=1}^{K} (a_k + c_k)$		$\sum_{k=1}^{K} n_k = N$

Obuchowski's test statistic for clustered matched-pair data:

$$X_O^{\,2} = \frac{(\hat{p}_1 - \hat{p}_2)^2}{\mathrm{var}(\hat{p}_1 - \hat{p}_2)_{\bar{p}}}$$

$X_O^{\,2}$ is asymptotically distributed as a chi-square with one degree of freedom under the null hypothesis.

Where

$$\mathrm{var}(\hat{p}_1 - \hat{p}_2)_{\bar{p}} = \mathrm{var}(\hat{p}_1)_{\bar{p}} + \mathrm{var}(\hat{p}_2)_{\bar{p}} - 2\,\mathrm{cov}(\hat{p}_1, \hat{p}_2)_{\bar{p}};$$
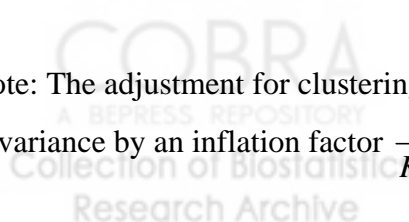
$$\bar{p} = (\hat{p}_1 + \hat{p}_2)/2;$$

$$\hat{p}_i = \frac{\sum\limits_{k=1}^{K} x_{ik}}{\sum\limits_{k=1}^{K} n_k} \quad \text{with (procedure i=1,2);}$$

$$\mathrm{var}(\hat{p}_i)_{\bar{p}} = \frac{K}{K-1}\left[\frac{\sum\limits_{k=1}^{K} (x_{ik} - n_k\,\bar{p})^2}{\left(\sum\limits_{k=1}^{K}(n_k)\right)^2} \right] \quad \text{with (procedure i=1,2);}$$

$$\mathrm{cov}(\hat{p}_i, \hat{p}_{i'})_{\bar{p}} = \frac{K}{K-1}\left[\frac{\sum\limits_{k=1}^{K} (x_{ik} - n_k\,\bar{p})\,(x_{i'k} - n_k\,\bar{p})}{\left(\sum\limits_{k=1}^{K} n_k\right)^2} \right] \quad \text{with (procedure i=1, i'=2).}$$

Note: The adjustment for clustering is taken into account by multiplying the variance and covariance by an inflation factor $\dfrac{K}{K-1}$.

Where we have a random sample of K clusters from a population, with n_k units in the k th cluster, k=1,...,K.

Each unit is given I treatments or tests.

Let x_{ik} denote the number of units in the k th cluster that respond to treatment i, i=1,...,I.

Let p_i denote the probability that a randomly selected unit from the population will respond to treatment i.

$\hat{p}_i$ is an estimator of p_i which is the overall proportion of responses to treatment.

Example:

Table 2. Specificity of PET versus SPECT (21 clusters; 51 matched-pairs).

Patient No. (cluster k)	Number of glands (n_k)	$x_{(PET)k}$	$x_{(SPECT)k}$
1	3	0 [0,0,0]	2 [0,1,1]
2	3	2 [1,1,0]	3 [1,1,1]
3	3	3 [1,1,1]	3 [1,1,1]
4	1	1 [1]	1 [1]
5	3	2 [1,1,0]	3 [1,1,1]
6	4	4 [1,1,1,1]	4 [1,1,1,1]
7	3	3 [1,1,1]	3 [1,1,1]
8	2	2 [1,1]	2 [1,1]
9	2	2 [1,1]	1 [1,0]
10	1	1 [1]	1 [1]
11	3	2 [1,1,0]	2 [1,1,0]
12	2	2 [1,1]	2 [1,1]
13	3	3 [1,1,1]	3 [1,1,1]
14	2	2 [1,1]	2 [1,1]
15	2	0 [0,0]	2 [1,1]
16	3	2 [1,1,0]	2 [1,1,0]
17	3	2 [1,1,0]	2 [1,1,0]
18	3	2 [1,1,0]	3 [1,1,1]
19	2	2 [1,1]	2 [1,1]
20	1	1 [1]	1 [1]
21	2	2 [1,1]	2 [1,1]

x_{ik} is the number of true negative test results for patient (cluster) k, test i.
The numbers in brackets correspond with the test results for each gland, where 1 indicates that the gland was called "true negative" for condition (success) and 0 indicates that the gland was called "false positive" (failure).
The PET test result listed first in the brackets corresponds with the SPECT test result listed first, etc. For example, for patient number 1 we have 3 matched-pairs: (0,0), (0,1), and (0,1).
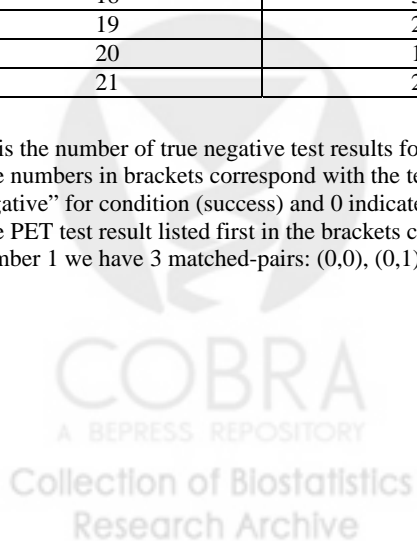
$$\sum_{k=1}^{K} n_k = 51 \qquad \sum_{k=1}^{K} x_{(PET)k} = 40 \qquad \sum_{k=1}^{K} x_{(SPECT)k} = 46$$

$$\hat{p}_{PET} = 40/51 = 0.784 \qquad \hat{p}_{SPECT} = 46/51 = 0.902$$

$$\text{var}(\hat{p}_{PET} - \hat{p}_{SPECT})\overline{p} = 0.00484$$

$$X_O^{\,2} = 2.88$$

p = 0.090   No rejection of the H_o

If the McNemar's Test is used without consideration of adjustment for clustered data, we get different results:

$$X_{Mc}^{\,2} = 4.5$$

p = 0.034   Rejection of the H_o

Durkalski's Method:

This is a simple method for the analysis of clustered matched-pair data that adjusts for multiple units within a cluster, yet avoids correlation assumptions among and within clusters and avoids distributional assumptions.

The null hypothesis is $H_0 : p_{10k} - p_{01k} = 0$ for all K clusters, where

$$p_{10k} = E\left[\frac{b_k}{n_k}\right] \quad \text{and} \quad p_{01k} = E\left[\frac{c_k}{n_k}\right]$$

A general form of the test statistic for **H₀:** $p_{10k} - p_{p01k} = 0$ for all K clusters, asymptotically distributed as a chi-square with one degree of freedom, is

$$X_G^{\,2} = \frac{(\hat{p}_{10} - \hat{p}_{01})^2}{\text{var}(\hat{p}_{10} - \hat{p}_{01})}$$

where $(\hat{p}_{10} - \hat{p}_{01}) = \dfrac{1}{K}\sum_{k=1}^{K}\dfrac{1}{n_k}(b_k - c_k)$

and
$$\mathrm{var}\left(\hat{p}_{10}-\hat{p}_{01}\right)=\frac{1}{K^{2}}\sum_{k=1}^{K}\left[\frac{1}{n_{k}}\left(b_{k}-c_{k}\right)-\left(\hat{p}_{10}-\hat{p}_{01}\right)\right]^{2}.$$

Under the null hypothesis is **H₀:** $p_{10k}-p_{p01k}=0$ for all K clusters,

$$\mathrm{var}\left(\hat{p}_{10}-\hat{p}_{01}\right)=\frac{1}{K^{2}}\sum_{k=1}^{K}\left[\frac{1}{n_{k}}\left(b_{k}-c_{k}\right)\right]^{2}.$$

Durkalski's test statistic is:

$$X_{V}^{2}=\frac{\left(\sum_{k=1}^{K}\frac{1}{n_{k}}\left(b_{k}-c_{k}\right)\right)^{2}}{\sum_{k=1}^{K}\left[\frac{1}{n_{k}}\left(b_{k}-c_{k}\right)\right]^{2}}.$$

Table 3. 2 x 2# Table Representation of the Specificity of PET versus SPECT (for use with the Durkalski Method).
 # Based on Table 1. McNemar's-like table for K clusters of matched-pair data.

Patient No. (cluster k)	Number of glands (n_k)	Matched- pair data	a_k	b_k	c_k	d_k	1/ n_k	b_k - c_k
1	3	(0,0) (0,1) (0,1)	0	0	2	1	1/3	-2
2	3	(1,1) (1,1) (0,1)	2	0	1	0	1/3	-1
3	3	(1,1) (1,1) (1,1)	3	0	0	0	1/3	0
4	1	(1,1)	1	0	0	0	1/1	0
5	3	(1,1) (1,1) (0,1)	2	0	1	0	1/3	-1
6	4	(1,1) (1,1) (1,1) (1,1)	4	0	0	0	1/4	0
7	3	(1,1) (1,1) (1,1)	3	0	0	0	1/3	0
8	2	(1,1) (1,1)	2	0	0	0	1/2	0
9	2	(1,1) (1,0)	1	1	0	0	1/2	1
10	1	(1,1)	1	0	0	0	1/1	0
11	3	(1,1) (1,1) (0,0)	2	0	0	1	1/3	0
12	2	(1,1) (1,1)	2	0	0	0	1/2	0
13	3	(1,1) (1,1) (1,1)	3	0	0	0	1/3	0
14	2	(1,1) (1,1)	2	0	0	0	1/2	0
15	2	(0,1) (0,1)	0	0	2	0	1/2	-2
16	3	(1,1) (1,1) (0,0)	2	0	0	1	1/3	0
17	3	(1,1) (1,1) (0,0)	2	0	0	1	1/3	0
18	3	(1,1) (1,1) (0,1)	2	0	1	0	1/3	-1
19	2	(1,1) (1,1)	2	0	0	0	1/2	0
20	1	(1,1)	1	0	0	0	1/1	0
21	2	(1,1) (1,1)	2	0	0	0	1/2	0

Table 4. Key to Interpretation of Matched-Pair Data.

PET	SPECT	
	Success	**Failure**
Success	$a_k = (1,1)$	$b_k = (1,0)$
Failure	$c_k = (0,1)$	$d_k = (0,0)$

Thus, $X_V^2 = 2.32$

and

$p = 0.128$ No rejection of the H_o.

Table 5. Comparison of Methods.

Method	Test Statistic	p-value	Rejection of H_0
Unadjusted	4.5	0.034	Yes
Obuchowski	2.88	0.090	No
Durkalski	2.32	0.128	No

***Confidence Interval for a Common Difference based on the Durkalski Method:***

It is useful to look at a non-null variance estimate if one is interested in calculating a confidence interval for the common difference $\hat{p}_{10} - \hat{p}_{01}$ (this assumes that the difference is constant across the K clusters).

If one assumes a common difference, then the variance estimate under the non-null condition, $H_A : p_{10k} - p_{01k} \neq 0$, is:

$$\operatorname{var}\left(\hat{p}_{10} - \hat{p}_{01}\right) = \frac{1}{K^2} \sum_{k=1}^{K} \left[\frac{1}{n_k}\left(b_k - c_k\right) - \left(\hat{p}_{10} - \hat{p}_{01}\right) \right]^2 .$$

Then, a 95 per cent confidence interval for the true difference $p_{10} - p_{01}$ is:

$$\left(\hat{p}_{10} - \hat{p}_{01}\right) \pm 1.96\sqrt{\left\{\operatorname{var}\left(\hat{p}_{10} - \hat{p}_{01}\right)\right\}} .$$
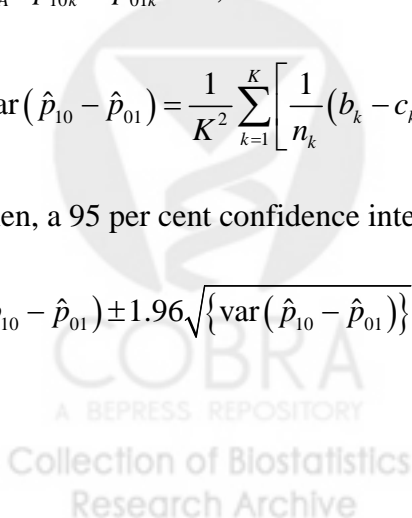
References:

[1] Obuchowski NA (1998). On the comparison of correlated proportions for clustered data. Statistics in Medicine 17, 1495-1507.

[2] Durkalski VL, Palesch YY, Lipsitz SR and Rust PF (2003). Analysis of clustered matched-pair data. Statistics in Medicine 22: 2417-2428.

[3] Conover WJ (1980). Practical nonparametric statistics, 2nd Edition, Wiley, NY.