Oyeka and Okeh, 2:2
http://dx.doi.org/10.4172/scientificreports.640

Open Access Scientific Reports

Research Article Open Access

S Coefficient of Concordance in Diagnostic Screening Tests

Oyeka ICA[1] and Okeh UM[2]*

[1]Department of Applied Statistics, Nnamdi Azikiwe University, Awka, Nigeria
[2]Department of Industrial Mathematics and Applied Statistics, Ebonyi State University, Abakaliki, Nigeria

Abstract

This paper proposes a measure of the strength of association, agreement or concordance between state of nature or condition in a population and test results in diagnostic screening tests. The proposed measure here termed "S Coefficient of Concordance" because it is a function of the sensitivity and specificity of the screening test, is normal between 0 and 1, assuming the value 0 when there is independence, complete discordance or disagreement and the value 1, when there is complete agreement or similarity between state of nature or condition and test results. Unlike the traditional odds ratio, the proposed S statistic in its specification and formulation properly ensures the non inclusion of the number of subjects in the sampled population who test negative even though they are actually positive in nature and the number of subjects who test positive even though they actually do not have the condition in nature, since these numbers are usually not known in diagnostic screening tests. The method develops the standard deviation of the proposed measure as well as test statistics that enables the testing of any desired hypothesis about not only the proposed S measure but also the sensitivity and specificity of the screening test. The method is illustrated with some sample data and the proposed 'S' measure is shown to be relatively more efficient and hence likely to be more powerful than the traditional odds ratio measure of association.

Keywords: Concordance; Coefficient; Diagnostic screening tests; Odds ratio; Sensitivity; Specificity

Introduction

Ordinarily in measuring the strength of association between two variables of classification either in cross sectional or longitudinal studies especially in medical research, the odds ratio, relative risk and other such measures rather than the Phi-coefficient are preferably used because unlike the later the former two measures are invariant under the three commonly used study methods [1,2]. However because the odds ratio and relative risk are not easy to clearly interpret and understand, some researchers prefer to use Berkson's simple difference or Shep's relative difference between rates as measures of association in medical research [1]. Unfortunately these two last measures are not invariant under the various designs. Also none of these measures of association theoretically has an upper bound to quickly indicate how large such measures could possibly be to indicate perfect disagreement or agreement between the variables of classification [3]. When used in the analysis of data obtained from diagnostic screening tests, a probably more serious problem with the traditional odds ratio and relative risk as measures of association is that they often include in their specifications and formulations the number of subjects testing negative among the subjects in the population known or believed to actually have the condition in nature and the number of subjects testing positives among those subjects known or believed not to actually have the condition in nature [4]. These are sample values that are in fact usually not readily known in diagnostic screening test results and ought not to be used in such analysis without prior modifications of the measures formulation. Thus the expressions for the traditional odds ratio, its standard deviation and significance test statistic are strictly speaking improperly specified and sometimes include unknown sample values.

When the prevalence rate of a condition in a population is known, then a measure of association between state of nature and test results in screening tests should ideally be based on true and false rates of the test that always factors in the prevalence rate. But prevalence rate of conditions in a population are not always known so that measures of association that include prevalence rates have limited applications [1,5].

In this paper we propose a measure of the strength of association or concordance between state of nature or condition and screening test results that does not require knowledge of the prevalence rate of the condition in the population. The proposed measure is based only on the sample data usually obtained in diagnostic screening tests, namely the total number of subjects studied consisting of the number of subjects drawn from the population known or believed to actually have the condition, the number drawn from the same population known or believed not to actually have the condition, the number testing positive among those subjects known to have the condition and the number of subjects testing negative among the subjects known not to have the condition. The proposed test statistic here termed the "S" coefficient of concordance because it is based on that is, it a function of only the sensitivity and specificity of the screening test, is normed to have values between 0 and 1 unlike the traditional odds ratios and relative risk measures of association which have no upper bounds.

Estimates of the standard error and test statistic appropriate for use with data from diagnostic screening tests are also proposed.

The Proposed Method

Suppose a medical researcher or clinician collects a random sample of $n_{.1}$ subjects from a population known or believed to have a certain condition in nature and another random sample of $n_{.2}$ subjects from the same population believed not to have the condition in nature giving a total random sample of $n=n..=n_{.1}+n_{.2}$ subjects to be studied. Research interest is in confirming through a diagnostic screening test whether or

***Corresponding author:** Okeh UM, Department of Industrial Mathematics and Applied Statistics, Ebonyi State University Abakaliki, Nigeria, E-mail: uzomaokey@ymail.com

Received January 16, 2013; **Published** February 28, 2013

Citation: Oyeka ICA, Okeh UM (2013) S Coefficient of Concordance in Diagnostic Screening Tests. 2: 640 doi:10.4172/scientificreports.640

not a randomly selected subject from the population has or does not have the condition of interest.

Let B be the event that a randomly selected subject from the population has the condition in nature and $\overline{B}$ be the event that the randomly selected subject does not have the condition in nature. Let A and $\overline{A}$ be respectively the events that a randomly selected subject from the population tests and does not test positives to the condition in the screening test. The results from such a screening test may be presented as in table 1.

To develop the proposed S measure of association or concordance we may let

$$u_{i1=}\begin{cases}1, \textit{if the ith sampled subject known or believed to have the condition in nature tests positive in the screening test}\\ 0, \textit{otherwise}\end{cases} \quad (1)$$

for i=1,2,...,$n_{.1}$

Let

$$\pi_1 = P(u_{i1} = 1) \quad (2)$$

Also Let

$$W_1 = \sum_{i=1}^{n_1} u_{i1} \quad (3)$$

Now

$$E(u_{i1}) = \pi_1; Var(u_{i1}) = \pi_1(1-\pi_1) \quad (4)$$

Also from Equations 3 and 4 we have that the expressed value of W_1 is

$$E(W_1) = \sum_{i=1}^{n_1} E(u_{i1}) = n_{.1}.\pi_1 \quad (5)$$

Similarly from Equation 4

$$Var(W_1) = n_{.1}.\pi_1(1-\pi_1) \quad (6)$$

Now π_1 is the probability that or the proportion of subjects testing positive among the sample of subjects known or believed to have the condition in nature, which is actually a measure of the sensitivity Se of the screening test. That is the proportion of subjects testing positive among those subjects in the population known or believed to have the condition in nature.

The sample estimate of π_1 is

$$\hat{\pi}_1 = \hat{Se} = \frac{f^+}{n_{.1}} = \frac{W_1}{n_{.1}} \quad (7)$$

Where f^+ is the total number of 1^s in u_{i1}, which is the number of subjects testing positive among the $n_{.1}$ subjects known or believed to have the condition in nature. The sample estimate of the variance of $\hat{\pi}_1$ is from Equations 6 and 7

$$Var(\hat{\pi}_1) = Var(\hat{Se}) = \frac{Var(W_1)}{n_{.1}^2} = \frac{\hat{\pi}_1(1-\hat{\pi}_1).}{n_{.1}} = \frac{\hat{Se}(1-\hat{Se})}{n_{.1}} \quad (8)$$

A researcher may for some reason wish to test the null hypothesis that the sensitivity of a test is some specified value $\pi_{10} = S_{e0}$: that is the null hypothesis

$$H_0 : \pi_1 = \pi_{10} (or\ Se = S_{e0}) \text{ versus } H_1 : \pi_1 \neq \pi_{10} (0 \leq \pi_{10} \leq 1) \quad (9)$$

The null hypothesis of Equation 9 is tested using the test statistic

$$\chi^2 = \frac{(W_1 - n_{.1}.\pi_{10})^2}{n_{.1}\hat{\pi}_1(1-\hat{\pi}_1)} = n_{.1}\frac{(\hat{\pi}_1 - \pi_{10})^2}{\hat{\pi}_1(1-\hat{\pi}_1)} \quad (10)$$

Which has approximately the chi-square distribution with 1 degree of freedom, the null hypothesis is rejected at the α level of significance if

$$\chi^2 \geq \chi^2_{1-\alpha;1} \quad (11)$$

Otherwise H_0 is accepted.

Now the estimated proportion of all samples subjects who have the condition in nature and also test positive to the condition is from Equation 7.

$$\hat{\pi}^+ = \frac{f^+}{n} = \frac{n_{.1}(f^+/n_{.1})}{n.} = \frac{n_{.1}.\hat{\pi}_1}{n} = \frac{n_{.1}.\hat{Se}}{n} \quad (12)$$

The corresponding estimated variance is

$$Var(\hat{\pi}^+) = \left(\frac{n_{.1}}{n}\right)^2 . Var(\hat{\pi}_1) = \frac{n_1\hat{\pi}_1(1-\hat{\pi}_1)}{n^2} = \frac{n_1\hat{Se}(1-\hat{Se})}{n^2} \quad (13)$$

Furthermore let

$$u_{i2} = \begin{cases}1, \textit{if the ith subject among the subjects known or believed not to have the condition in nature tests negative in the screening test}\\ 0, \textit{otherwise}\end{cases} \quad (14)$$

for i=1,2,…,$n_{.1}$

Let

$$\pi_2 = P(u_{i2} = 1)$$

Also let

$$W_2 = \sum_{i=1}^{n_2} u_{i2} \quad (16)$$

Now

$$E(u_{i2}) = \pi_2; Var(u_{i2}) = \pi_2(1-\pi_2) \quad (17)$$

Also

$$E(W_2) = \sum_{i=1}^{n_2} E(u_{i2}) = n_{.2}.\pi_2 \quad (18)$$

and

$$Var(W_2) = n_{.2}.\pi_2(1-\pi_2) \quad (19)$$

Now π_2 is the proportion of subjects testing negative among the subjects known or believed not to have the condition in nature. Note that π_2 is actually the specificity Sp of the test; that is, the proportion of subjects testing negative among those subjects in the population known or believed not to have condition in nature. Its sample estimate is

$$\hat{\pi}_2 = \hat{Sp} = \frac{f^-}{n_2} = \frac{W_2}{n_2} \quad (20)$$

The corresponding estimated variance is

$$Var(\hat{\pi}_2) = Var(\hat{Sp}) = \frac{Var(W_2)}{n_2^2} = \frac{\hat{\pi}_2(1-\hat{\pi}_2).}{n_2} = \frac{\hat{Sp}(1-\hat{Sp})}{n_2} \quad (21)$$

Again for some reasons a researcher may wish to test the null hypothesis that sensitivity of a test is some specified value $\pi_{20} = S_{p0}$ or the null hypothesis

$$H_0 : \pi_2 = \pi_{20} (or\ Sp - S_{p0}) \text{ versus } H_1 : \pi_2 \neq \pi_{20} (0 \leq \pi_{20} \leq 1) \quad (22)$$

$$\chi^2 = \frac{(W_2 - n_{.2}.\pi_{20})^2}{n_2\hat{\pi}_2(1-\hat{\pi}_2)} = n_{.2}\frac{(\hat{\pi}_2 - \pi_{20})^2}{\hat{\pi}_2(1-\hat{\pi}_2)} \quad (23)$$

which has approximately the chi-square distribution with 1 degree of freedom, the null hypothesis is rejected at the α level of significance if Equation 11 is satisfied otherwise H_0 is accepted.

Now the proportion of all sampled subjects in the population who are known or believed not to have the condition and also test negative in the screening test is

$$\hat{\pi}^{-} = \frac{f^{-}}{n} = \frac{n_2(f^{-}/n_2)}{n.} = \frac{n_2.\hat{\pi}_2}{n} = \frac{n_2.\hat{S}p}{n} \quad (24)$$

With estimated variance

$$Var(\hat{\pi}^{-}) = \left(\frac{n_2}{n}\right)^2 . Var(\hat{\pi}_2) = \frac{n_2\hat{\pi}_2(1-\hat{\pi}_2)}{n^2} = \frac{n_2\hat{S}p(1-\hat{S}p)}{n^2} \quad (25)$$

The expected value of $\hat{\pi}^{+}$, namely $\hat{\pi}^{+} = \frac{n_1.Se}{n.}$ is the proportion of the population who are known or believed to have a condition in nature and also test positive in a the screening test. Similarly the expected value of $\hat{\pi}^{-}$ namely $\hat{\pi}^{-} = \frac{n_2.Se}{n.}$, is the proportion of the population who are known or believed not to have the condition in nature and also test negative in the screening test. Their sum, namely

$$\pi = S = \frac{n_1.Se}{n} + \frac{n_2.Sp}{n} = \frac{n_1.Se + n_2.Sp}{n} \quad (26)$$

Is the proportion of all the subjects in the population who either are known or believed to have the condition in nature and also test positive or known or believed not to have the condition in nature and also test negative in the screening test. The larger the value of π=*S*, the smaller is its complement namely the proportion of subjects in the population who either are known or believed to have the condition in nature but test negative or known or believed not have the condition in nature but positive in the screening test, or vice versa.

In effect this relationship implies that the higher or larger the value of π=*S*, the stronger is the association, agreement or concordance between state of nature or condition and test results, the lower or smaller the value, the weaker is the agreement or concordance between state of nature and test results.

Now since π=*S* is a probability, it can only assume values that are between 0 and 1 inclusively. In other words, $0 \le S \le 1$.The proposed measure of association π=*S*, agreement or concordance assumes the value 0 when there is independence, complete discordance or lack of any agreement between state of nature or condition and test results. In this case there is no association whatsoever between state of nature and test results, so that knowledge of a subjects test result is of no use in predicting or determining whether or not the subject has or does not have the condition in nature. If *S*=1, then there is perfect positive agreement or concordance between state of nature or condition and test results, so that knowledge of a subjects test result would enable accurate prediction of the subjects true condition in nature. Normally S assumes intermediate values between 0 and 1.

The sample estimate of S, the proposed measure of concordance, agreement or strength of association between state of nature or condition and test results in diagnostic screening tests is therefore $\hat{\pi} = \hat{S} = \hat{\pi}^{+} + \hat{\pi}^{-}$, which is a weighted mean of the sensitivity and specificity of the screening test, estimated from Equation 26 as

$$\hat{\pi} = \hat{S} = \hat{\pi}^{+} + \hat{\pi}^{-} = \frac{n_1\hat{S}e + n_2\hat{S}p}{n} \quad (27)$$

Note that S is invariant under the three common study methods used for medical research and assumes all possible values between 0 and 1 inclusively. A researcher may wish to test the null hypothesis of no association that is of independence H_0: *S*=0 between state of nature or condition and test results. However, a more general hypothesis would be

$$H_0 : \pi = \pi_0 (or\ S = S_0) \text{ versus } H_1 : \pi \ne \pi_0 (or\ S \ne S_0)(0 \le \pi_0, S_0 \le 1) \quad (28)$$

Now to estimate the variance of $\hat{\pi} = \hat{S}$, we note that by their specifications in Equations 1 and 14, u_{i1} and u_{i2} are uncorrelated so that from Equations 13 and 25 we have that

$$\frac{Var(\hat{\pi})}{n^2} = \frac{Var(\hat{S})}{n^2} = \frac{Var(\hat{\pi}^{+}) + Var(\hat{\pi}^{-})}{n^2} = \frac{n_1\hat{\pi}_1(1-\hat{\pi}_1) + n_2\hat{\pi}_2(1-\hat{\pi}_2)}{n^2} = \frac{n_1\hat{S}e(1-\hat{S}e) + n_2\hat{S}p(1-\hat{S}p)}{n^2} \quad (29)$$

Hence the test statistic for the null hypothesis of Equation 28 is

$$\chi^2 = \frac{(\hat{\pi} - \hat{\pi}_0)^2}{Var(\hat{\pi})} = \frac{(\hat{S} - \hat{S}_0)^2}{Var(\hat{S})} = \frac{(n_1\hat{S}e + n_2\hat{S}p - n\hat{S}e)^2}{n_1\hat{S}e(1-\hat{S}e) + n_2\hat{S}p(1-\hat{S}p)} \quad (30)$$

which has approximately, the chi-square distribution with one degree of freedom for sufficiently large 'n' and may be used to test the null hypothesis of Equation 28. The null hypothesis is rejected at the α level of significance if Equation 11 is satisfied, otherwise H_0 is accepted.

As noted above the usual practice has often been to use the traditional odds ratio or relative risk to assess the strength of association between state of nature or condition and test results in diagnostic screening tests. A problem with this approach however is that the number of subjects testing positive among those known to be free of the condition and the number of subjects testing, negative among those subjects known to have the condition of interest are not usually known and hence also are their derivatives $n_{1.}$ and $n_{2.}$, the total number of subjects that would either test positive or negative to the test. Hence all calculations including measures of association, their precision and test statistics based on these unknown values are fundamentally faulty and without modifications would yield misleading and unsatisfactory results.

The proposed S measure of concordance and the associated test statistic which are functions of the sensitivity and specificity of the screening test are not encumbered by these problems. Furthermore the proposed method enables the testing of any desired hypotheses concerning sensitivity and specificity of the test which provide useful information and an additional advantage over and above other existing methods.

Illustrative Example

A research scientist collected a random sample of n.1=28 subjects from a certain community known or believed to have breast cancer and also a random sample of n_2=49 subjects from the same community, known or believed not to have breast cancer. Research interest is to confirm through a diagnostic screening test whether each of the

Screening Test	(Condition) Present	(Condition) Absent	Total($n_{i.}$)
Result	B	$\bar{B}$	
Positive (*A*)	n_{11}=f^{+}	n_{12}	$n_{1.}$
Negative($\bar{A}$)	n_{21}	n_{22}=f^{-}	$n_{2.}$
Total($n_{.j}$)	$n_{.1}$	$n_{.2}$	n (=n..)

Table 1: Formation for the presentation of result of a diagnostic screening test. State of Nature (Condition).

Clinical diagnosis	Histologic Diagnosis(state of nature)		
Test Results	Breast Cancer present (*B*)	Breast Cancer Absent ($\bar{B}$)	Total ($n_{i.}$)
Positive (*A*)	n_{11}=23=f^{+}	n_{12}=7	$n_{1.}$=30
Negative ($\bar{A}$)	n_{21}=5	n_{22}=42=f^{-}	$n_{2.}$=47
Total ($n_{.j}$)	$n_{.1}$=28	$n_{.2}$=49	n..=n=77

Table 2: Results of the Screening Test for Breast Cancer in a Certain Community.

sampled subjects has or does not have breast cancer. The results of the screening test are presented in table 2.

We here use the data of table 2 to illustrate the proposed method.

Results and Discussions

Now from table 2 and Equations 7 and 20 we have that the estimated sensitivity and specificity of the test are respectively,

$$\hat{Se} = \hat{\pi}_1 = \frac{23}{28} = 0.821 \text{ and } \hat{Sp} = \hat{\pi}_2 = \frac{42}{49} = 0.857$$

These values show that the screening test is quite sensitive and specific. The corresponding estimated variances are from Equation 8 and 21 respectively.

$$Var(\hat{Se}) = Var(\hat{\pi}_1) = \frac{(0.821)(1-0.821)}{28} = 0.005$$

and $$Var(\hat{Sp}) = Var(\hat{\pi}_2) = \frac{(0.857)(1-0.857)}{49} = 0.003$$

Now from Equation 12 we have that the estimated proportion of all sample subjects who have the condition in nature and also test positive is

$$\hat{\pi}^+ = \frac{28(0.821)}{77} = 0.299$$

With estimated variance Equation 13 given as

$$Var(\hat{\pi}^+) = \frac{(28)(0.821)(1-0.821)}{(77)^2} = 0.0007$$

Similarly from Equation 24 we have that the estimated proportion of the sampled population who are known or believed not to have the condition and also test negative in the screening test is

$$\hat{\pi}^- = \frac{(49)(0.858)}{77} = 0.545$$

Whose estimated variance is from Equation 25

$$Var(\hat{\pi}^-) = \frac{(49)(0.857)(1-0.857)}{(77)^2} = 0.001$$

Therefore from Equation 27, we have that the sample estimate of the proposed measure of association or coefficient of concordance between state of nature (existence of breast cancer in the population) and screening test results (presence of breast cancer as revealed by the test) is

$$\hat{S} = \hat{\pi}^+ + \hat{\pi}^- = 0.299 + 0.545 = 0.844 \text{ or } 84.4\%$$

This estimated value of $\hat{S} = 0.844$ *or* 84.4% suggests that there is a high level of agreement or concordance between actual existence of breast cancer in the population and the screening test results.

The sample variance of this estimate is from Equation 29

$$Var(\hat{\pi}) = Var(\hat{S}) = 0.0007 + 0.001 = 0.0017$$

Hence the test statistic for the null hypotheses of no association between state of nature and test results, that is, the null hypothesis of Equation 28 with H_0=S_0=0 is from Equation 30

$$\chi^2 = \frac{(0.844)^2}{0.0017} = 419.021 (P - value = 0.0000)$$

which with one degree of freedom is highly statistically significant indicating a high level of association or concordance between the existence of breast cancer in the sampled population and screening test results.

Now if we have used the traditional odds ratio *W*=o as a measure of association we would have from table 2 that the estimated sample odds ratio is

$$W = o = \frac{n_{11}n_{22}}{n_{12}n_{21}} = \frac{(23)(42)}{(7)(5)} = 27.600$$

The estimated standard deviation is given as

$$se(o) = \sqrt[o]{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}} = \sqrt[27.6]{\frac{1}{23} + \frac{1}{7} + \frac{1}{5} + \frac{1}{42}} = 17.664$$

The corresponding chi-square Test statistic is

$$\chi^2 = \frac{n(n_{11}n_{22} - n_{12}n_{21})^2}{n_1 n_2 n_{.1} n_{.2}} = \frac{\left((77)(23)(42) - (7)(5)\right)^2}{(30)(47)(28)(49)} = 34.504 (P - value = 0.0000)$$

which with 1 degree of freedom is also highly statistically significant. Note however that the estimated value of the proposed coefficient of concordance or measure of association is $\hat{S} = 0.844$ which has by specification is always at most 1.0, its upper limit compared with the estimated odds ratio O=27.600 which theoretically has no upper bound. Also the standard error of $\hat{S}$ is only $\hat{S} = 0.041$ which is much less than the estimated standard error of O of se(O)=17.664, showing that the proposed measure is much more precise and efficient than the traditional odds ratio. Also note that the proposed method and the traditional odds ratio method here both lead to a rejection of the null hypothesis of independence or no association, the relative sizes of the calculated chi-square values show that at least for the present example the odds ratio method is likely to lead to an acceptance of a false null hypothesis (Type II error) more often and hence is likely to be less powerful than the proposed S method.

Summary and Conclusion

We have in this paper proposed a measure of association, agreement or concordance, between state of nature or condition in a population and test results in diagnostic screening tests and it is formulated to be a function of the sensitivity and specificity of the screening test. Unlike the traditional odds ratio method which by specification has no upper limit and assumes the value 1 if there is no association, the proposed statistics, the coefficient of concordance, S is normed to assumed values between 0 and 1 inclusively with 0 indicating perfect discordant or independence and 1 indicating perfect agreement or concordance between state of nature and screening test result. Also the specification and formulation of the proposed measure of association S unlike the traditional odds ratio intrinsically and structurally does not include in its specification and formulation the number of subjects in the sampled population testing negative among those subjects who are known or believed to have the condition in nature and the number of the subjects testing positive among those who are known or believed not to have the condition in nature. These sample values are usually not known in diagnostic screening tests and cannot be properly and validly used without modification in analysis of data from diagnostic screening tests.

The standard error of the proposed statistic and a test statistic for its significance are developed. It is shown using sample data that the proposed test statistic is more efficient and powerful than the traditional odds ratio method.

References

1. Fleiss JL (1981) Statistical Method for Rates and proportion. (2ndedn), John Wiley & Sons, New York, USA.
2. Raslich MA. Markert RJ, Stutes SA (2007) Selecting and interpreting diagnostic tests. Biochemia Medica 17: 151-161.
3. Agresti A (2007) An Introduction to Categorical Data Analysis. (2ndedn), John Wiley & Sons, New Jersey, USA.
4. Bossuyt PMM (2006) Clinical evaluation of medical tests: still a long road to go. Biochemia Medica 16: 103-106.
5. Greenberg RS, Daniels SR, Dana Flanders W, Eley JW, Boring JR (2001) Medical Epidemiology, Lange-McGraw-Hill, London, UK.