Science Journal of Mathematics & Statistics
ISSN: 2276-6324
http://www.sjpub.org/sjms.html

Published By Science Journal Publication

Research Article **Volume 2013 (2013), Article ID sjms-106, 4 Pages, doi: 10.7237/sjms/106**

Sensitivity and Specificity Dependent Measure of Association in Diagnostic Screening Tests

Oyeka ICA [1], Igwebuike V. Onyiaorah [2], Adaora Onyiaorah [3], Chilota Efobi [4] and Okeh UM [5].

[1]Department of Applied Statistics, Nnamdi Azikiwe University, Awka Nigeria.
[2]Department of Histopathology, Nnamdi Azikiwe University Teaching Hospital Nnewi Anambra State, Nigeria.
[3]Department of Opthalmology, Enugu State University Teaching Hospital Park lane Enugu State, Nigeria.
[4]Deparment of Haematology, University of Port Harcourt Teaching Hospital,Port Harcourt, Rivers State Nigeria.
[5]Department of Industrial Mathematics and Applied Statistics, Ebonyi State University Abakaliki, Nigeria.

Accepted 2nd June, 2013

Abstract

This paper proposes and presents a measure of the strength of association between test results and state of nature or condition in a population constructed using only sensitivity and specificity of diagnostic screening tests. The proposed measure which always has between -1 and 1 inclusively enables the researcher to determine not only if an association exists between test results and condition, but if such an association exists, whether it is positive and direct or negative and indirect thereby giving the measure an advantage over the traditional odds ratio method. Estimates of the standard error and test statistic for the proposed measure are provided. Results using the proposed measure are shown to be easier to interpret and understand than those obtained using the traditional odds ratio approach. Furthermore, using sample data, the proposed measure is shown to be at least as efficient and hence as powerful as the traditional odds ratio.

Keywords: traditional odds ratio, relative risk, sensitivity, specificity, state of nature or condition

Introduction

The traditional odds ratio and relative risk are often used as measures of the strength of association between a predisposing or antecedent factor and condition in controlled comparative studies because these two measures are invariant under the three commonly used study methods to generate the data, namely the cross-sectional, prospective and retrospective study designs (Fleiss, 1981).However in diagnostic screening tests these measures cannot strictly speaking properly and validly be used(Greenberg et al,2001). This is because in these tests the number of subjects testing positive among subjects known or believed not to have a condition and the number of subjects testing negative among subjects known or believed to have a condition in nature usually are not known and hence the total number of subjects testing positive and negative are not also completely known (Pepe,2003; Linn,2004).These values which are contained in the expressions used in the estimation of the traditional odds ratio and relative risk, their associated standard errors and test statistics for significance can therefore not be properly directly used in calculations. The false rates and the proportion of the population expected to test positive should ideally be factored –in and reflected in measures used to assess association in diagnostic screening tests (Pepe, 2003; Fliess, 1981; Linn, 2004).Unfortunately the utility of these three indices is seriously limited by the fact that their formulations contain prevalence rates of the conditions in a population of interest, values that often are known for many conditions. This paper proposes to develop a statistical measure of the strength of association between diagnostic screening test results and state of nature or condition in a population that depends on only the sensitivity and specificity of screening tests that are estimable using only the observed sample values. The proposed measure does not also require knowledge of the prevalence rate of a condition in the population before its estimation.

The Proposed Method

Suppose a research scientist or clinician collects a random sample of $n_{.1}$ subjects known or believed to actually have a certain condition in a population and also collects a second random sample of $n_{.2}$ subjects from the same population known or believed not to actually have the condition in nature, giving a total random sample of size $n=n_{..}=n_{.1}+n_{.2}$ subjects to be studied. Research interest is to determine through a diagnostic screening test whether or not each of the sampled subjects actually tests positive or negative to the condition.

Let B be the event that a randomly selected subject actually has a condition in nature and $\bar{B}$ be the event that the randomly selected subject does not have the condition in nature. Also let A *and* $\bar{A}$ be respectively the events that the randomly selected subject tests and does not test positive to the condition in the screening test. The results of such a screening test may be presented as in Table 1

Table 1: Format for Presentation of Results in Diagnostic Screening Test

	State of Nature or Condition		
Test Result	**Present (B)**	**Absent ($\bar{B}$)**	**Total ($n_{i.}$)**
Positive (A)	n_{11}	n_{12}	$N_{1.}$
Absent($\bar{A}$)	n_{21}	n_{22}	$N_{2.}$
Total ($n_{.j}$)	$n_{.1}$	$n_{.2}$	$n_{..}(=n)$

In Table 1, out of $n_{.1}$ subjects known or believed to have a condition in nature n_{11} subjects test positive and n_{21} test negative. Similarly, out of $n_{.2}$ subjects known or believed not to have a condition in nature n_{12} test positive and n_{22} test negative .Of the $n=n_{..}$ Subjects sampled, $n_{1.}$ Subjects test positive while $n_{2.}$ Subjects test negative. However as noted above, in diagnostic screening test results only n_{11} and n_{22} subjects which are often of primary interest to the researcher are usually known. The values n_{12}, the number of subjects testing positive among those known or believed not to have the condition and n_{21}, the number of subjects testing negative among those known or believed to have the condition usually are not known. Hence the marginal totals $n_{1.}$ And $n_{2.}$ Usually are not completely known so that these unknown values may not properly and validly be used in calculations. The present proposed measure of association is based on the expectation that if a diagnostic screening test is a good one, then the sum of the proportion of subjects testing positive among the population subjects known or believed to have the condition in nature would be much larger than the sum of the proportion of subjects testing positive among the population of subjects known or behaved not to have the condition and the proportion of subjects testing negative among the population of subjects known or believed to have the condition in nature. Similarly if the screening test is a poor one then one would expect a converse result, that is, the sum of proportion of subjects testing positive among the population of subjects known or believed not to have a condition and the proportion of subjects testing negative among the population of subjects known or believed to have the condition would be larger than the sum of the proportion of subjects testing positive among the population of subjects known or believed to have a condition and the proportion of subjects testing negative among the population of subjects known or believed not to have the condition in nature. If there is no association between screening test results and state of nature or condition, these two sums would be equal and the difference between the two would be zero.

Now using conditional probabilities of events, the proportion of subjects testing positive among the population of subjects known or

believed to have a condition in nature which is actually the sensitivity, Se of the screening test (Linn, 2004;Baron 2001) is

The proportion of subjects testing negative among the population of subjects known or believed not to have a condition in nature which is also the specificity, Sp of the screening test is

$$P(A|B) = \frac{P(AB)}{P(B)} = Se \qquad 1$$

$$P(\bar{A}|\bar{B}) = \frac{P(\bar{A}\bar{B})}{P(\bar{B})} = Sp \qquad 2$$

Also the proportion of subjects testing negative among the population of subjects known or believed to have a condition in nature is

$$P(\bar{A}|B) = 1 - P(A|B) = 1 - Se \qquad 3$$

And the proportion of subjects testing positive among the population of subjects known or believed not to have a condition in nature is

$$P\left(A|\bar{B}\right) = 1 - P\left(\bar{A}|\bar{B}\right) = 1 - Sp \qquad 4$$

Hence the difference between the proportions of subjects testing positive among subjects known or believed to have a condition or testing negative among subjects known or believed not to have the condition and the proportions of subjects testing positive among subjects known or believed not to have a condition or testing negative among subjects known or believed to have a condition is

$$® = \left(Se + Sp\right) - \left((1 - Se) + (1 - Sp)\right) = 2\left(Se + Sp - 1\right) \qquad 5$$

To use Equations 1-5 to develop the proposed measure of association we may let

$$u_i = \begin{cases} 1, \text{if the ith sampled and screened subject either tests positive} \\ \quad \text{and is known or believed to actually have the condition or tests negative} \\ \quad \text{and is known or believed not to actually have the condition in nature} \\ -1, \text{if the ith sampled and screened subject either tests positive and is known} \\ \quad \text{or believed not to have the condition or tests negative and is known or} \\ \quad \text{believed to actually have the condition in nature} \end{cases} \qquad 6$$

$$for\ i = 1, 2, .., n$$

$$Let$$

$$®^{+} = P(u_i = 1); ®^{-} = P(u_i = -1) \qquad 7$$

$$where\ ®^{+} + ®^{-} = 1 \qquad 8$$

$$Define$$

$$W = \sum_{i=1}^{n} u_i \qquad 9$$

Now

$$E(u_i) = ®^{+} - ®^{-}; Var(u_i) = ®^{+} + ®^{-} - \left(®^{+} - ®^{-}\right)^2 = 1 - \left(®^{+} - ®^{-}\right)^2 \qquad 10$$

$$Also$$

$$E(W) = \sum_{i=1}^{n} E(u_i) = n(®^{+} - ®^{-}) \qquad 11$$

And

$$Var(W) = \sum_{i=1}^{n} Var(u_i) = n\left(®^{+} - ®^{-} - (®^{+} - ®^{-})^2\right) \qquad 12$$

Now $®^{+}$ is the probability that a randomly selected subject either is positive and is actually positive in nature or tests negative and is actually negative in nature; while $®^{-}$ is the probability that a randomly selected subject either test positive and is negative in nature or test negative and is positive in nature. Their sample estimates using the frequencies in Table 1 i.e are respectively

$$\hat{®}^{+} = \frac{f^{++} + f^{--}}{n} = \frac{n_{.1}\left(f^{++}/n_{.1}\right) + n_{.2}\left(f^{--}/n_{.2}\right)}{n} = \frac{n_{.1}\hat{Se} + n_{.2}\hat{Sp}}{n} \qquad 13$$

where $\hat{Se}$ *and* $\hat{Sp}$ are respectively the sample estimates of sensitivity and specificity of the screening test and

$$\hat{®}^{-} = \frac{f^{+-} + f^{-+}}{n} = \frac{(n_{.2} - f^{--}) + (n_{.1} - f^{++})}{n} = \frac{\left(n_{.2} - n_{.2}\left(f^{--}/n_{.2}\right)\right) + \left(n_{.1} - n_{.1}\left(f^{++}/n_{.1}\right)\right)}{n} = \frac{n_{.1}(1 - \hat{Se}) + n_{.2}(1 - \hat{Sp})}{n} \qquad 14$$

Hence the sample estimate of the difference between these probabilities or proportions is

$$\hat{®} = \hat{®}^{+} - \hat{®}^{-} = \frac{\left(f^{++} + f^{--}\right) - \left(f^{+-} + f^{-+}\right)}{n} = \frac{W}{n} \qquad 15$$

With variance estimated using Equation 12 as

$$Var\left(\hat{®}\right) = Var\left(\hat{®}^{+} - \hat{®}^{-}\right) = \frac{Var(W)}{n^2} = \frac{1 - \left(\hat{®}^{+} - \hat{®}^{-}\right)^2}{n} \qquad 16$$

When expressed in terms of estimated sensitivity and specificity of the test Equation 15 becomes

$$\hat{®} = \hat{®}^{+} - \hat{®}^{-} = \frac{n_{.1}(2\hat{Se} - 1) + n_{.2}(2\hat{Sp} - 1)}{n} \qquad 17$$

With corresponding variance obtained from Equation 16 as

$$Var(\hat{®}) = Var\left(\hat{®}^{+} - \hat{®}^{-}\right) = \frac{1 - \left(\frac{n_{.1}(2\hat{Se} - 1) + n_{.2}(2\hat{Sp} - 1)}{n}\right)}{n} \qquad 18$$

The hypothesis that is usually of research interest is whether or not the diagnostic screening test is a good one, that is whether an association exists between test results and condition. As noted above, if there is no association between screening test results and state of nature or condition then the proportion $® = ®^{+} - ®,^{-}$ that is the difference between the sum of the proportion of subjects who either test positive among the population of subjects known or believed to have a condition or the proportion of subjects testing negative among subjects in the population known or believed not to have the condition and the sum of the proportion of subjects who either test positive among the population known or believed not to have the condition on the proportion of subjects who test negative among the population of subjects known or believed to have the condition in nature would be expected to be thus zero symbolically a null hypothesis of interest would be

$$H_0 : ® = ®^{+} - ®^{-} \le 0 \ versus\ H_1 : ® = ®^{+} - ®^{-} > 0, \qquad 19$$

The null hypothesis of Equation 19 may be tested using the test statistic

$$¡^2 = \frac{\hat{®}^2}{Var(\hat{®})} = \frac{\left(\hat{®}^{+} - \hat{®}^{-}\right)^2}{Var\left(\hat{®}^{+} - \hat{®}^{-}\right)} = \frac{\left(n_{.1}(2\hat{Se} - 1) + n_{.2}(2\hat{Sp} - 1)\right)^2}{n\left(1 - \left(\frac{n_{.1}(2\hat{Se} - 1) + n_{.2}(2\hat{Sp} - 1)}{n}\right)\right)^2} = \frac{\left(n_{.1}(2\hat{Se} - 1) + n_{.2}(2\hat{Sp} - 1)\right)^2}{n - \frac{\left(n_{.1}(2\hat{Se} - 1) + n_{.2}(2\hat{Sp} - 1)\right)^2}{n}} \qquad 20$$

Which has approximately the chi-square distribution with 1 degree of

Freedom for sufficiently large 'n'. The null hypothesis H_0 is rejected at the ~ level of significance if

$$¡^2 \ge ¡^2_{1-\sim;1} \qquad 21$$

Otherwise H_0 is accepted.

Illustrative Example

We here use the following data to illustrate the proposed method. A clinician collected a random sample of 98 subjects from a certain

How to Cite this Article: Oyeka ICA, Igwebuike V. Onyiaorah, Adaora Onyiaorah, Chilota Efobi and Okeh UM, "Sensitivity and Specificity Dependent Measure of Association in Diagnostic Screening Tests" Science Journal Of Mathematics and Statistics, Volume 2013 (2013), Article ID sjms-106, 4 Pages, doi: 10.7237/sjms/106

population; 12 of whom are suspected to have prostrate cancer and 86 of whom are believed not to have the disease. The clinician's interest is to confirm through a diagnosis screening test whether or not each of the sampled subjects is actually prostrate cancer positive or negative. The results of the screening test are presented in Table 2.

Table 2: Result of Prostrate Cancer Screening Test

Clinical diagnosis	**Present (B)**	**Absent($\bar{B}$)**	
Prostrate Cancer	$n_{11}=f^{++}=4$	$n_{12}=f^{+-}=4$	$n_{1.}=6$
Positive(A)	$n_{21}=f^{-+}=4$	$n_{22}=f^{--}=4$	$n_{2.}=92$
Negative ($\bar{A}$)	$n_{.1}=12$	$n_{.2}=86$	$n_{..}=n=98$

Now from Table 2 we have that the sample estimate of the sensitivity and specificity of the test are respectively

$$\hat{S}e = \frac{f^{++}}{n_{.1}} = \frac{n\left(f^{++}/n\right)}{n_{.1}} = \frac{98\left(4/98\right)}{12} = (8.167)(0.041) = 0.335$$

and

$$\hat{S}p = \frac{f^{--}}{n_{.2}} = \frac{n\left(f^{--}/n\right)}{n_{.2}} = \frac{98\left(84/98\right)}{86} = (1.140)(0.857) = 0.977.$$

These results show that the screening test is low in sensitivity but has high specificity.

Now from Equations 13 and 14 the sample estimates of $\hat{R}^{+}$ *and* $\hat{R}^{-}$ are respectively

$$\hat{R}^{+} = \frac{12(0.335) + 86(0.977)}{98} = 0.898$$

and

$$\hat{R}^{-} = \frac{12(1-0.335) + 86(1-0.977)}{98} = 0.102$$

Hence from Equations 15 and 17 we have that

$$\hat{R} = \hat{R}^{+} - \hat{R}^{-} = 0.898 - 0.102 = 0.796$$

With estimated variance obtained from Equations 16 and 18 as

$$Var(\hat{R}) = Var(\hat{R}^{+} - \hat{R}^{-}) = \frac{1-(0.796)^{2}}{98} = 0.004$$

Hence the test statistic of no association between screening test results and state of nature or condition (Prostrate Cancer) of Equation 19 is obtained from Equation 20 as

$$\chi^{2} = \frac{(0.796)^{2}}{0.004} = \frac{0.634}{0.004} = 158.500(P-value = 0.0000)$$

Which with 1 degree of freedom is highly statistically significant indicating a strong degree of association between screening test results and state of nature or condition (presence of Prostrate cancer in the population).Also since $\hat{R} = \hat{R}^{+} - \hat{R}^{-} = 0.796$ is positive, the association is positive and direct.

It would be instructive to compare the present results with what would have been obtained if we had used the traditional odds ratio to analyze the data of Table 2 in spite of odds ratios short comings as already pointed out above when used in the analysis of screening test results. The sample estimate of the traditional odds ratio for the data of Table 2 is

$$O = \frac{n_{11}n_{22}}{n_{12}n_{21}} = \frac{(4)(84)}{(2)(8)} = 21.00$$

This means that for every one subject who has prostrate cancer among those tested and erroneously informed that they are free of the disease 21 subjects among those tested and found to have prostrate cancer would be expected to be correctly so informed. This is probably more difficult to understand than the simple information conveyed by the simple difference in rates, $\hat{R} = \hat{R}^{+} - \hat{R}^{-} = 0.796$, namely, that the proportion of subjects testing positive among subjects who have Prostrate cancer or testing negative among subjects who do not have prostrate cancer is 79.6 percent higher than the proportion of subjects testing positive among subjects who do not have prostrate cancer or testing negative among subjects who have the disease. In other words, the number of subjects who test positive among those who actually have prostrate cancer or negative among subjects who do not have the disease is about 79.6 percent more than the number of subjects who test positive among subjects known not to have the disease or who test negative among subjects known or believed to actually have prostrate cancer. The standard error of the estimated odds ratio is

$$Se(O) = O\sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}} = 21.00\sqrt{\frac{1}{4} + \frac{1}{2} + \frac{1}{8} + \frac{1}{84}}$$
$$= (21.00)(0.942) = 19.782$$

This measure of the error of O namely 19.782 is clearly much larger than the error of only $Se(\hat{R}) = \sqrt{0.004} = 0.063$ of the estimated value of $\hat{R}$ for our sample data. The chi-square test statistic for the significance of O is

$$\chi^{2} = \frac{n(n_{11}n_{22} - n_{12}n_{21})^{2}}{n_{1.}n_{2.}n_{.1}n_{.2}} = \frac{98(94)(84) - (2)(8))^{2}}{(6)(92)(12)(86)} = 17.616(P-value = 0.0000)$$

Which is also statistically significant again leading to a rejection of the null hypothesis of no association. However, although the proposed method and the traditional odds ratio approach here both lead to a rejection of the null hypothesis, the relative sizes of the calculated chi-square values suggest that the traditional odds ratio method is less efficient and likely to lead to an acceptance of a false null hypothesis (Type II Error) more frequently and hence is likely to be less powerful than the proposed method.

Summary and Conclusion

We have in this paper proposed, developed and presented a statistical method for measuring the strength of association between test results and state of nature or condition in a population expressed to a diagnostic screening test. The proposed measure is based on only the sensitivity and specificity of the screening test which are independent of the population of interest and estimated using only observed sample values.

The proposed measure which always lies between -1 and 1 inclusively can be used to establish whether an association is strong and direct, strong and indirect or zero estimates of the standard error and test statistic for the significance of the proposed measure are provided. The proposed measure of association is shown to be easier to interpret and explain than the traditional odds ratio, and the sample data used suggest that the measure is at least as efficient and powerful as the traditional odds ratio.

Reference

How to Cite this Article: Oyeka ICA, Igwebuike V. Onyiaorah, Adaora Onyiaorah, Chilota Efobi and Okeh UM, "Sensitivity and Specificity Dependent Measure of Association in Diagnostic Screening Tests" Science Journal Of Mathematics and Statistics, Volume 2013 (2013), Article ID sjms-106, 4 Pages, doi: 10.7237/sjms/106

1. Baron JA. 2001. Clinical epidemiology, in Teaching Epidemiology eds. Olsen J; Saracci R; and Trichopoulos D; Oxford:Oxford University Press, 237-249.

2. Fleiss JL. 1981. Statistical Method for Rates and proportions (2nd ed), New York:John Wiley & Sons.

3. Greenberg RS, Daniels SR, Flanders WD, Eley JW and Borin g JR. 2001. Medical Epidemiology, London: Lange-McGraw- Hill.

4. Pepe MS. 2003. The Statistical Evaluation of Medical Tests for Classification and Prediction. Oxford statistical series 28, Oxford University Press, U.K.

5. Linn, Shai 2004. A new Conceptual Approach to teaching the interpretation of clinical tests. Journal of statistics education 12(3). www.amstat.org/publications/jse/v12n3/linn.html.

How to Cite this Article: Oyeka ICA, Igwebuike V. Onyiaorah, Adaora Onyiaorah, Chilota Efobi and Okeh UM, "Sensitivity and Specificity Dependent Measure of Association in Diagnostic Screening Tests" Science Journal Of Mathematics and Statistics, Volume 2013 (2013), Article ID sjms-106, 4 Pages, doi: 10.7237/sjms/106