European Journal of Applied Sciences 4 (2): 90-97, 2012
ISSN 2079-2077

DOI: 10.5829/idosi.ejas.2012.4.2.65132

Understanding and Applications of Test Characteristics and Basic Inferential Statistics in Hypothesis Testing

Ajay S. Singh and Micah B. Masuku

Department of Agricultural Economics and Management, University of Swaziland, Swaziland

Abstract: Scientific researches and diagnostics tools in medical and applied sciences have an important role to play in the health care system as well as agricultural production and development of a nation. In view of radical change in the research spectrum, the scenario is becoming difficult and interesting for the researchers and associated scholars. The statisticians design the experiment, analyze the data and interpret the facts with the help of traditional statistical techniques and statistical inference helps to draw the conclusions in scientific manner. Characteristic for diagnostic test provides the idea to physician in true assessment of clinical disease and statistical inference provides the idea or guide to scientists in the testing of research hypothesis and their interpretation. Sensitivity, specificity, positive predictive value and negative predictive value are collectively known as test characteristics. It is more important ways to express the usefulness of diagnostic tests. It is also more important to understanding of sensitivity, specificity, positive predictive value, negative predictive value and their significant applications and interpretation in applied sciences. Generally, test characteristics guide the clinician in assessment of disease entities. In a similar manner statistical inference guide the researcher in the testing of research hypothesis and interpretation. It is necessary to understand the basics of test characteristics and hypothesis testing to gain appreciation. These test characteristics and statistical inferences are more useful in medical and agricultural sciences (animal science, plant pathology, etc.). In this article, we discussed the basic understanding to calculate sensitivity, specificity, positive predictive value and negative predictive value and their significant interpretation and also discussed the basic statistical inferential techniques. We have discussed the importance of these measures and provided how we should use these measures in our day-to-day applied research.

Key words: Sensitivity · Specificity · Predictive value · Hypothesis Testing · Confidence Interval

INTRODUCTION

Advanced applied scientific researches have experienced a dramatic change in knowledge and an exponential increase in technology. A lot of these technical developments involve applied researches. Sensitivity and specificity are statistical measures of the performance of a binomial classification test, also known in statistics as classification function. Sensitivity are also called recall rate in some fields and measures the proportion of actual positives which are correctly identified as such (e.g. the percentage of sick people who are correctly diagnosed). Specificity measures the proportion of negatives which are correctly identified (e.g. the percentage of healthy people who are not correctly diagnosed). These two measures are closely related to the concepts of statistical type I and type II errors. A perfect predictor would be described as 100% sensitivity (i.e. predict all people from the sick group as sick) and 100% specificity (i.e. not predict anyone from the healthy group as sick), theoretically any predictor will possess a minimum error bound known as the Bayes error rate.

Binomial classification is the task of defining the members of a given set of objects into two strata on the basis of whether they have some property or not. Some typical binomial classifications are given below.

Medical testing to identify if a patient has certain disease or not (the binomial classification property is the disease).

Corresponding Author: Ajay S. Singh, Department of Agricultural Economics and Management, University of Swaziland, Swaziland.

Quality control in factories; i.e. deciding if a new product is good enough to be sold, or if it should be discarded (the binomial classification property is being good enough).

Deciding whether a page or an article should be in the result set of a search or not (the binomial classification property is the relevance of the article - typically the presence of a certain word in it).

In this article we try to explain the test characteristics i.e. sensitivity, specificity, predictivity and their scientific application in the applied research especially in medical and agricultural sciences. Sensitivity, specificity, positive predictive value and negative predictive value are known as test characteristics and these are important ways to express the usefulness of diagnostic tests. The test outcome can be positive (predicting that the person has the disease) or negative (predicting that the person does not have the disease). The test results for each subject may or may not match the subject's actual status. In this article we also try to explain the basic statistical inference in terms of hypothesis, types of error and their probabilities, p values, confidence interval etc.

Definitions and Concepts (Validity, Sensitivity and Specificity)

Validity: It is the extent to which a test measures what it is supposed to measure; in other words, it is the accuracy of the test. Validity is measured by sensitivity and specificity.

Different fields in epidemiology have different levels of validity. One way to assess the validity of findings is the ratio of false-positives to false-negatives. The validity of a study is dependent on the degree of systematic error. Validity is usually separated into two components [1-4].

There are two types of validity, known as internal and external validity. Internal validity is dependent on the amount of error in measurements, including exposure of disease and the associations between the variables. Good internal validity implies a lack of error in measurement and suggests that inferences may be drawn. External validity pertains to the process of generalizing the findings of the study to the population from which the sample was drawn. This requires an understanding of which conditions are relevant to the generalization [1-4].

Sensitivity: Sensitivity relates to the test's ability to identify positive results. Sensitivity is the true positive rate [5-7]. In other word with the example of the medical test used to identify a disease. The sensitivity of a test is the proportion of people who have the disease who test positive for it.

Sensitivity = Number of True Positives / (Number of True positives + Number of False Negatives)

This indicates that the probability of positive test.
For clarification:

True positive: Ill people correctly diagnosed as ill
False positive: Healthy people incorrectly identified as ill
True negative: Healthy people correctly identified as healthy
False negative: Sick people incorrectly identified as healthy.

If a test has high sensitivity then a negative result would suggest the absence of disease. For example, a sensitivity of 100% means that the test recognizes all actual positives i.e. all sick people are recognized as being ill. Thus, in contrast to a high specificity test, negative results in a high sensitivity test are used to rule out the disease.

In non-medical contexts, sensitivity is sometimes called recall. Sensitivity is not the same as the precision or positive predictive value (ratio of true positives to combined true and false positives), which is as much a statement about the proportion of actual positives in the population being tested as it is about the test.

The calculation of sensitivity does not take into account indeterminate test results. If a test cannot be repeated, the options are to exclude indeterminate samples from analysis or, alternatively, indeterminate samples can be treated as false negatives. A test with a high sensitivity has a low type II error rate [5-8].

Specificity: Specificity relates to the ability of the test to identify negative results. Specificity is the "true negative rate [5-7].

In the example of the medical test used to identify a disease. The specificity of a test is defined as the proportion of patients who do not have the disease and will test negative for it.

Specificity = Number of True negatives / (Number of True Negatives + Number of False Positives)

This specificity indicates the probability of a negative test. If a test has high specificity, a positive result from the test means a high probability of the presence of a disease. In other word, Specificity is the "true negative rate," equivalent to [D/B +D]. Positive predictive value (PPV) is the proportion of people with a positive test result who actually have the disease [A/A+ B]; Negative predictive value (NPV) is the proportion of those with a negative result who do not have the disease [D/C+ D].

Sensitivity and specificity are fixed for a particular type of test. Positive Predictive Value and Negative Predictive Value for a particular type of test depend upon the prevalence of a disease in a population.

For example, though current screening tests for blood cancer have high sensitivity and specificity, the low prevalence of blood cancer in the general population cannot justify universal screening since the majority of positive tests would be falsely positive (i.e. low PPV).

Sensitivity and specificity are inversely proportional, meaning that as the sensitivity increases, the specificity decreases and vice versa [5-8].

Concept of Gold Standard: The gold standard is the best single test that is considered the current preferred method of diagnosing a particular disease. All other methods of diagnosing disease, including any new test, need to be compared against this 'gold' standard. The gold standard is different for different diseases. The gold standard for any disease X may be considered outdated or inadequate, but any new test designed to replace the gold standard has to be initially validated against the gold standard. If the new test is indeed better, there are ways to prove that; following which the new test may become the gold standard. Common example of gold standards include the use of electro cardio graphic changes plus cardiac enzyme levels to diagnose acute myocardial infaraction, or pulmonary angiography to diagnose pulmonary embolism and it will be assumed that results obtained by gold standard tests are always correct [9-11].

Calculations:

Sensitivity = A / A+C
= A (true positive) / A+C (True Positive + False Negative)
= Probability of being test positive when disease present.
Specificity = D / B+D
= D (True Negative) / B+D (False Positive + True Negative)
= Probability of being test negative when disease absent.
Prevalence: The proportion of the population with disease
= 100 X [(TP+FN)/N]
Sackett [9] discussed the likelihood ratio to characterize diagnostic test.
Positive Likelihood Ratio Test = Sensitivity/ (1-Specificity)
Negative Likelihood Ratio = (1-Sensitivity)/Specificity
Sample Calculation

Suppose a patient with anemia and a serum ferritin of 60 mmol/L and come across a systematic review of serum ferritin as a diagnostic test for iron deficiency anemia, with the results summarized in the Table 2 below.

The sensitivity and specificity are calculated as follows:

Sensitivity = A / (A+C)
= 731/809 = 90%
Specificity = D / (B+D)
= 1500/1770 = 85%

These results show that 90% of the patients with iron deficiency anemia have a positive test result (serum ferritin <65 mmol/L), while 85% of patients who do not have the disorder test negative (serum ferritin ≥ 65 mmol/L). Since the specificity is high greater than 80% and patient has a Positive test result, can rule in the diagnosis of iron deficiency anemia.

In the other results, Harvey assessed the use of plasma D-dimer levels for diagnosing deep venous thrombosis (DVT) in patients hospitalized for stroke rehabilitation are summarized in the table below. The presence or absence of DVT was determined by positive or negative venous duplex ultrasound results [13].

The sensitivity and specificity are calculated as follows:

Sensitivity = A /(A+C) = 13/14 = 92.8%
Specificity = D/(B+D) = 72/91 = 79.1%

Definitions and Concepts of Predictive Value

Positive Predictive Value: It is the percentage of patients with a positive test who actually have the disease. In a 2 x 2 table (Table 1), cell 'a' is 'true positives' and cell 'b' is

Table: 1 2 x 2 table (For diagnostic test results)

	Disease		
Test Results	Present (+)	Absent (-)	Totals
Positive (+)	A or TP	B or FP	A+ B
Negative (-)	C or FN	D or TN	C+ D
Totals	A + C	B + D	A + B + C + D

Table 2: Distribution of patient with Iron Deficiency Anemia and level of Serum Ferritin

		Target Disorder (Iron Deficiency Anemia)		
Diagnostic Test (Serum Ferritin)		Present	Absent	Total
	+ (<65 mmol/L)	A 731	B 270	A+B 1001
	- (≥65 mmol/L)	C 78	D 1500	C+D 1578
Total		A+C 809	B+D 1770	A+B+C+D 2579

Source: Guyatt [12]

Table 3: Distribution of patient with deep venous thrombosis and level of Plasma D-dimer level

		Target Disorder (DVT)		
Diagnostic Test (PlasmaD-dimer level)		Present	Absent	Total
	+ (>1591 mmol/L)	A 13	B 19	A+B 32
	- (≤ 1591 mmol/L)	C 01	D 72	C+D 73
Total		A+C 14	B+D 91	A+B+C+D 105

'false positives.' In real life situation, we do the new test first and we do not have results of 'gold standard' available. We want to know how this new test is doing. PPV tells us about this – how many of test positives are true positives; and if this number is higher (as close to 100 as possible), then it suggests that this new test is doing as good as gold standard.

PPV: = A / A+B
= A (True Positive) / A+B (True Positive + False Positive)
= Probability (Patient having disease when test is positive)

Example: Sensitivity and Specificity provided to calculate positive predictive value.

PPV = A (True Positive) / A+B (True Positive + False Positive)

Negative Predictive Value (NPV): It is the percentage of patients with a negative test who do not have the disease. In 2 x 2 table (Table 1), cell 'D' is 'true negatives' and cell 'C' is 'False Negatives.' NPV tells us how many of test negatives are true negatives; and if this number is higher (should be close to 100), then it suggests that this new test is doing as good as gold standard.

NPV: = D / C+D
= D (True Negative) / C+D (False Negative + True Negative)
= Probability (patient not having disease when test is negative)

Example: Sensitivity and specificity provided in Table 1 to calculate negative predictive value.

NPV = A (True Negatives) / C+D (False Negative + True Negative) Positive and negative predictive values are directly related to the prevalence of the disease in the population. Assuming all other factors remain constant, the PPV will increase with increasing prevalence; and NPV decreases with increase in prevalence. This is illustrated by the following example.

As the disease prevalence increases, the positive predictive value also increases

Understanding of Predictive Value: In block/cell A, mark those in whom the test in question correctly diagnosed the disease (as determined by the gold standard). In other words, the test is positive, as is the gold standard. These are the true positives (TP) [10, 11].

In block/cell B, mark those who have positive results for the test in question but do not have disease according to the 'gold standard test.' The newer test has wrongly diagnosed the disease: These are false positives (FP) [10, 11].

In block/cell C, mark those who have disease on the 'gold standard test' but have negative results with the test in question. The test has wrongly labeled a diseased person as normal: These are false negatives (FN) [10, 11].

In block/cell D, mark those who have no disease as determined by the 'gold standard test' and are also negative with the newer test. These are true negatives (TN) [10, 11].

Statistical Inference: Hypothesis Testing: Inferential statistics or statistical inference includes the testing of hypothesis which is essential and important parts of research investigations. In traditional statistical hypothesis testing, the statistician starts with a null hypothesis and an alternative hypothesis, performs an experiment and then decides whether to reject the null hypothesis in favour of the alternative. In other words hypothesis is a numerical statement of about the parameter [14].

The first step in hypothesis is to state the null hypothesis H_o which follows logically from alternative hypothesis H_1 [14-16]. Alternative hypothesis define the research statement in positive terms [15]. Acceptance or rejection of null hypothesis based on our statistical testing parametric or non parametric methodologies [14, 16, 17]. If null hypothesis H_0 is accepted, then H_1 must be rejected and vice versa due to the hypothesis are mutually exclusive. If H_0 is accepted, this concludes that no statistical differences exist and if any differences in groups or observations due to only chance or due to sampling fluctuations. In other hand, if H_0 is rejected or H_1 is accepted this indicates that a significant difference exits and the differences are not only due to chance or sampling fluctuations.

Statistical Error in Hypothesis Testing: There are two types of error or incorrect conclusions are possible in hypothesis. Tables 1 shows that possibilities in which the statistical test falsely indicates that differences exists significantly between the two or more groups and also analogously to a wrong positive results. Rejection of null hypothesis H_0 when it is true is called as Type I error and acceptance of null hypothesis H_0 when it is false and it is known as Type II error and Type ii error is more harmful then Type I error [14-17].

The probability of Type I error is known as level of significance (α) and the probability of Type II error is known as the power of the test β or (1- α) [14, 15, 17]. By convention, statistical significance is generally accepted if the probability of making type I error is less than 0.05, which is commonly denoted as $p < 0.05$ [17, 18]. The probability of type ii error is more difficult to derive than probability of type I error, actually it is not one single probability value. The probability of type II error (β) is often ignored by researcher [19]. The probability of type I error (α) and probability of type II error (β) are inter related. As α arbitrarily decreased, β is increased. Similarly, α is increased, β is decreased [14, 15].

Statistical Power: Statistically power indicates mathematically the probability of not making a type II error. Statistical Power is defined as (1-β) [14, 15]. β indicates the probability of making II error and if sample size increases, power increases [16, 18].

Power is analogous to sensitivity in hypothesis testing. Sensitivity indicates the probability that the diagnostic test can detect disease when it present. Power indicates the probability that the statistical test can detect significant differences, when in fact such differences are truly exists.

P-values: p value is the probability to observe effects as big as those seen in the study if there is really no difference between the groups or treatments. The reasoning of hypothesis testing and p values is convoluted. p values helps to assessing whether this apparent effect is likely to be actual or could just by chance or sampling fluctuation. p values gives the magnitude of difference present between population. In calculation of p values, first assume that no true difference between the two groups/treatments. p values allows the assessment of findings are significantly different or not statistically different. If p value is small, the findings are unlikely to have arisen by chance or sampling fluctuation, reject the null hypothesis. If p is large, the observed difference is plausibly a chance finding, we do not reject the null hypothesis. By convention, p value of less than 5% is considered small or significant. Sometimes p value is less than 1% or 0.01, called as highly significant [16, 20].

Confidence Interval: Confidence interval, like p-values, provides a guide to helps the interpretation of research findings in the light of the probability. Confidence interval describes the different information from that arising in hypothesis test. Confidence interval provides a range about the observed effect size. The formal definition of confidence interval is a range of values for a variable of interest constructed so that this range has a specified probability is called the confidence level and the end

points of confidence interval are called the confidence limits [21]. By conventional, confidence interval at the 95% this corresponds to hypothesis testing with p-values, with a cut off for p is less than 0.05. A different and potentially more useful approach to assessing the role of chance has come to the fore confidence intervals [16, 20, 22].

Understanding of Predictive Values with Type of Errors: A positive or statistically significant result is one which rejects the null hypothesis. Doing this when the null hypothesis is in fact true - a false positive - is a type I error; doing this when the null hypothesis is false results in a true positive. A negative or not statistically significant result is one which does not reject the null hypothesis. Doing this when the null hypothesis is in fact false - a false negative - is a type II error; doing this when the null hypothesis is true results in a true negative.

To measure the performance of a medical test, the concepts sensitivity and specificity are used. Say we test some people for the presence of a disease. Some of these people have the disease and our test says they are positive. They are called true positives (TP). Some have the disease, but the test claims they don't. They are called false negatives (FN). Some don't have the disease and the test says they don't - true negatives (TN). Finally, we might have healthy people who have a positive test result - false positives (FP). Thus, the number of true positives, false negatives, true negatives and false positives add up to 100% of the set.

Applications: Dubey *et al.* [23] reported sensitivity and specificity of various serological tests for the detection of toxoplasma gondii infection in naturally infected soces. Montasser *et al.* [24] used sensitivity and specificity for the efficacy of serological tests. Naithani *et al.* [25] studied the patients with suspected pulmonary embolism multi detector CTA- CTV has higher diagnostic sensitivity than does CTA alone with similar specificity. Prabhu [26] studied and reported sensitivity and specificity in the rapid detection of Pnca mutations in Pyrazinamide resistant mycobacterium tuberculosis isolates. Yuen and Hughes [27] presented the examples of sensitivity and specificity based on data of Jones [28]. Duraisamy [29] discussed the cost effective methods for detecting cervical cancer with adequate sensitivity and specificity and reported that ideal screening test is one that is minimally invasive, easy to perform, acceptable to the subject, cost effective and easily invasive state with adequate sensitivity and specificity. Rao [30] discussed the meaning of sensitivity, specificity and predictive values evidenced based medicine. Yuen [31], Metz [32] and Hughes *et al.* [33] also used in fruitful studies. Parikh [10] discussed the basic knowledge to calculate the sensitivity and specificity and used these results for our patients in day to day life. Hui and Walter [34] also studied the consistent maximum likelihood estimates of the test sensitivity and specificity for two or tests.

Scott *et al*. [35] studied the correlation factor for estimating the agricultural injuries through ambulance report methods and calculate the sensitivity and specificity. Daniel *et al*. [36] reported sensitivity and specificity of the complement fixation test for the detection of cattle persistently infected with anaplasma marginale. Yuen and Mila [37] used Bayesian approaches and predictive system in plant pathology [37]. Darwish *et al*. [38] studied the evaluation of PCR assay for detection of cow's milk in water buffalo's milk and after the evaluation of sensitivity and specificity of the primers was established, its applicability on milk samples.

CONCLUSION

An understanding and interpretation of diagnostic tests facilitates an understanding of hypothesis testing. A test result may be a true positive, true negative, false positive or false negative. Sensitivity and specificity are the characteristics of the diagnostic test and prevalence of disease or event is a determinant of the predictive value. Similarly, hypothesis in testing, type I error occurs with probability α and type II error with probability β or $1-\alpha$. Statistical power or Power $1-\beta$ is analogous to the sensitivity of a diagnostic test. As sensitivity and specificity are not predictive, therefore power is also not predictive. Prevalence of disease or events affects the predictive value of a positive test result and predictive value affects the statistical significant test. In this paper we provided the basic knowledge to calculate sensitivity, specificity, positive predictive value and negative predictive value. More importantly, we have discussed these measures and provided how we should use these measures in our day-to-day clinical practice and research and also have illustrated how to calculate sensitivity and specificity, while combining two tests and how to use the results for our patients in day-to-day practice.

REFERENCES

1. Rothman, K., 2002. Epidemiology. An introduction, Oxford University Press.

2. Rothman, K., G. Sander and L. Timothy, 2008. Modern Epidemiology, 3rd Edition, Lippincott Williams and Wilkins.
3. Ioannidis, J.P.A., R. Tarone and J.K. McLaughlin, 2011. The False-positive to False-negative Ratio in Epidemiologic Studies. Epidemiology, 22(4): 450-456.
4. Singh, A.S., 2012. Statistical Research design and basic techniques in epidemiological research, European Jr. of Applied Sciences, 4(1): 27-35.
5. Sacket, D.L., P. Tugwell and G.H. Guyatt, 1991. Clinical Epidemiology: A basic science for clinical medicines, 2nd edition, Bostan: Little Brown.
6. Alfredo Morabia, 2004. A history of epidemiologic methods and concepts. Birkhäuser.
7. Ray, M.M., 2010. Introduction to Epidemiology. Jones and Bartlett Learning.
8. Haynes, R.B., D.L. Dacket, G.H. Guyatt and P. Tugwell, 2006. Clinical epidemiology, how to do clinical practice research, 3rd edition, Lippincott and Wilkins, Philadelphia.
9. Sackett, EL., R.B. Haynes and P. Tugwell, 1985. Clinical epidemiology: a basic science for clinical medicine, Little Brown and Company, Boston.
10. Parikh, R., A.M. Parikh, G.S. Sekhar and R. Thomas, 2008. Ind. J. Ophthalmology, 56(1): 45-50.
11. Durham, T.A. and J.R. Turner, 2007. Introduction to statistics in pharmaceutical clinical trials, Pharmaceutical Press, London.
12. Guyatt, G.H., 1992. Laboratory diagnosis of iron deficiency anemia: Overview, J. Gen. Intern., 7: 145-53.
13. Harvey, R.L., E.J. Roth, P.R. Yarnold, J.R. Durham and G. David, 1996. Deep vein thrombosis in stroke: the use of plasma D-dimer level as a screeningtest in the rehabilitation setting, Stroke, 27: 1516-1520.
14. Hopkins, K.D. and G.V. Glass, 1978. Basic statistics for the behavioral sciences, Englewood Cliffs, New Jersey, Prentice Hall, USA.
15. Keppel, G., 1978. Design and analysis. A researcher's handbook. Englewood Cliffs, New Jersey, Prentice Hall. USA.
16. Gupta, S.C. and V.K. Kapoor, 1970. Fundamental of mathematical statistics, SC Publication, New Delhi, India.
17. Elenbaas, R.M., J.K. Elenbass and P.G. Cuddy, 1983. Evaluating the medical literature Part II: Statistical Analysis, Ann. Emerg. Med., 12: 610-613.
18. Sokal, R.R. and F.I. Rohlf, 1981. Biometry (ed. 2) New York, WH Freeman and Co.
19. Freeman, J.A., T.C. Chalmers, H. Smith *et al*., 1978. The importance of beta, the type II error and sample size in the design and interpretation of randomized clinical trial. New Engl. J. Med., 299: 690-694.
20. Rao, N.S.N., 1985. Elements of Health Statsitics, First edition, R. Publication, Varanasi, India.
21. Last, J.M., 1988. A dictionary of epidemiology, Oxford: International J. Epidemiology.
22. Gardner, M.J. and D.G. Altman, 1986. Confidence intervals rather than p values: estimation rather than hypothesis testing, BMJ, 292: 746-750.
23. Dubey, J.P., P. Thulliez, R.M. Weigel, C.D. Andrews, P. Linda and E.C. Powell, 1995. Sensitivity and specificity of various serologic tests for detection of Toxoplasma gondii infection in naturally infected sows. American Jr. Vet. Research, 56(8): 1030-36.
24. Montasser, A.M., M.M. Affi, E.M. El-Bayoumy, U.M. Abdul-Rouf and H.A. Mohammad, 2011. Efficiency of Serological Tests for Detection of Brucellosis in Ruminant at South Provinces of Egypt, Global Veterinaria, 6(2): 156-161.
25. Naithani, R., J. Chandra, V. Singh and V. Kumar, 2006. Life threatening exacerbation in idiopathic pulmonary hemosiderosis salvaged by cyclophosphamide infusion, The Indian J. Chest Disease and Allied Sciences, 48: 286-289.
26. Prabhu, S.S., M. Mutharaj and J. Muralidhar, 2009. Rapid detection of PnCa mutations in Pyrazinamide resistant mycobacterium tuberculosis clinical isolates, European J. App. Sc., 1(2): 20-25.
27. Yuen, J.E. and G. Hughes, 2002. Bayesian analysis po plant pathology, Plant pathology, 51(4): 407-412.
28. Jones, D.R., 1994. Evaluation of fungicides for control of eyespot disease and yield loss relationships in winter wheat. Plant Pathology, 43(5): 831-846.
29. Duraisami, K., K.S. Jaganathan and J.C. Bose, 2011. Advances in Biological Research, 5(4): 226-32.
30. Rao, G., 2004. Defining the terms of evidence medicine The problem with Ps. J. Family Practice, 53(3).
31. Yuen, J., E. Twengstrom and R. Sigvald, 1996. Calibration and verification of risk algorithms using logistic regression, European J. Plant Pathology, 102(9): 847-854.
32. Metz, C.E., 1978. Basic principles of RoC analysis, Seminars in Nuclear Medicine, 8(4): 283-298.
33. Hughes, G., Mc., N. Roberts and F.J. Burnett, 1999. Decision making and diagnosis in disease management, Plant Pathology, 48(2): 147-153.

34. Hui, S.L. and S.D. Walter, 1980. Estimating the error rates of diagnostic tests, Biometrics, 36: 167-71.
35. Scott, E.E., R.G. Earle, N. Krupa and P. Jenkins, 2011. A correction factor for estimating statewide agricultural injuries from ambulance reports, Ann. Epidemiology, 21(10): 767-72.
36. Daniel, S.B., Susana Torioni De Echaide, P.K. Donald, G.H. Steven, G. and Terry F. Mc Elwain, 2001. Sensitivity and specificity with compliment fixation test for detection of cattle persistently infection with anaplasma marginale, Jr. Vet. Diag. Invest., 13(79).
37. Yuen, J. and A. Mila, 2004. Are Bayesian approaches useful in plant pathology, Frontis, 8: 95-103.
38. Darwish, S.F., A.A. Hanaa and A.S. Amin, 2009. World Applied Sciences Journal, 7(4): 461-467.