pISSN 2394-6032 | eISSN 2394-6040

DOI: http://dx.doi.org/10.18203/2394-6040.ijcmph20181673

Review Article

Set theoretic approach to the concept of net sensitivity and net specificity in screening test results

Bansi Badan Mukhopadhyay, Himadri Bhattacharjya*

Department of Community Medicine, Agartala Government Medical College, Agartala, Tripura, India

Received: 22 February 2018
Accepted: 21 March 2018

***Correspondence:**
Dr. Himadri Bhattacharjya,
E-mail: hbhattacharjya@rediffmail.com

ABSTRACT

Net sensitivity and net specificity have been reviewed from set theoretic approach. With a basic knowledge of set theory one can estimate the net sensitivity and specificity in an easy way in both sequential and simultaneous screening tests. Union, intersection and complementary operations of set theory have been adopted to find out the solutions.

Keywords: Net sensitivity, Net specificity, True positive, True negative, Universal set, Union, Intersection, Complementary operations of set theory

INTRODUCTION

The subject of net sensitivity and specificity has been discussed by many authors mostly in descriptive ways for studying the validity of sensitivity tests.[1,2] The present article is an attempt to discuss the concept of set theoretic approach and its practical application in a lucid and user friendly manner. The concepts of union, intersection and complementation have been used for studying net sensitivity and specificity in two types of screening tests viz. sequential and simultaneous. To describe the theme of the subject, basic knowledge of set theory is essential.

Using Venn diagram one may see that in case of union operation if A and B are two sets in S, then AUB is also a set in S. This means the occurrence of A or B or both. In other words, AUB means the occurrence of at least one of the sets A and B. Thus AUB=A+B - A∩B (Figure 1).

Again using Venn diagram one can see that in case of intersection operation if A and B are two sets in S, then A∩B is also a set in S. This means the joint occurrence of A and B and is a set of common elements of A and B (Figure 2).

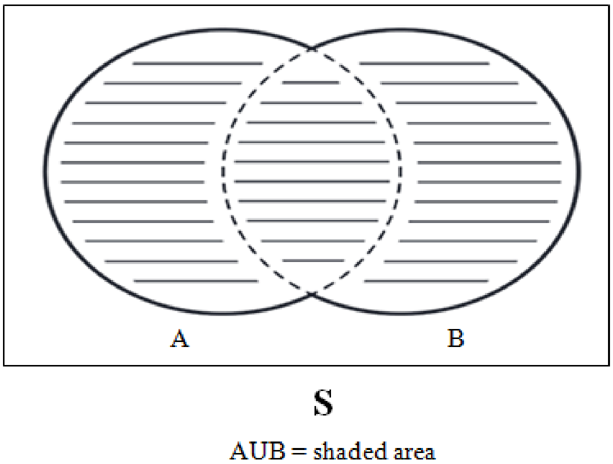
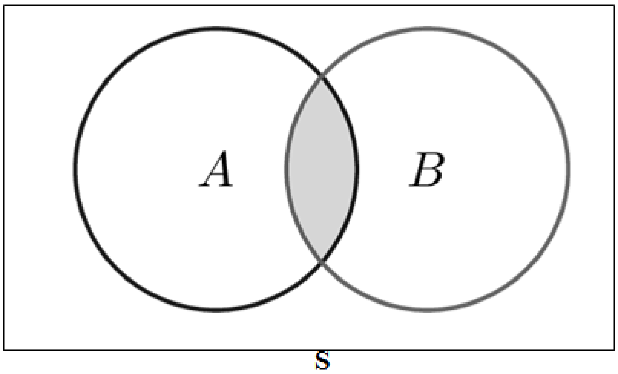


Figure 1: Venn diagram showing union operation.

Likewise, in case of difference operation if A and B are two sets in S, then A - B is the set which means the occurrence of A along with non-occurrence of B. A–B= $A∩B^C$ and B–A= $B∩A^C$.



$A \cap B$ = shaded area.

Figure 2: Venn diagram showing intersection operation.

Net sensitivity and specificity

In two types of screening tests viz. sequential and simultaneous, net sensitivity and specificity can be measured in the following ways.

Sequential screening test

In sequential screening test less expensive, less invasive and less efficacious test is usually performed at the first stage and those who tested positive are further tested with a superior test and the true positives are determined. Thus the false positive results are minimized. Here the results of test 2 are dependent on the results of test 1.

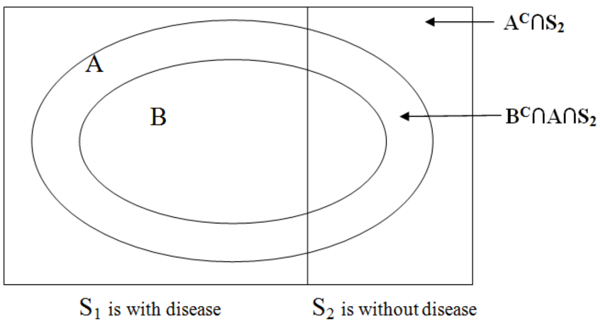


Figure 3: Venn diagram showing areas for calculation of net sensitivity and net specificity in sequential screening tests.

For clearer perspective, we use the following notations: S=Universal set of population under study. S_1=Set of those who have the disease under study determined on the basis of prevalence of the disease and $n(S_1)$ denotes their number. S_2=Set of those who do not have the disease under study. $n(S_1)+n(S_2)=n(S)$. A= the set of positives by test1 and n(A) denotes their number. Obviously, $A \cap S_1$=true positives, $A^C \cap S_2$=true negatives. $A \cap S_2$=false positive by test 1, $A^C \cap S_1$=false negative by test 1. B=set of positives by the superior test 2 among test positives of test 1. $(A \cap B \cap S_1)$=true positive by both the tests in sequential testing. Thus net sensitivity={True positive in both tests ÷ those who have the disease}×100={n $(A \cap B \cap S_1)$ ÷ n (S_1)}×100 (Figure 3).

Since in sequential screening test true positives are reduced at the second stage, it results in loss in net sensitivity. Hence $A - B = A \cap B^C$=Set of false positive results by test1 and $B^C \cap S_2$=Set of true negatives in test 2. Those who tested negative by test1 are not further tested by test 2 and among the test positive of test 1, further false positives are separated, resulting in some additional true negatives among the test positives of test 1. Thus $A \cap B^C \cap S_2$=true negative by test 2 among the test positives by test 1. Thus the total true negative results after two stage sequential screening test is: $n((A^C \cap S_2) U (A \cap S_2 \cap B^C))$=n $(A^C \cap S_2)$+n $(A \cap B^C \cap S_2)$ (since these are disjoint sets). Therefore net specificity={n $(A^C \cap S_2)$+ n $(A \cap B^C \cap S_2)$} ÷ {n (S_2)×100. Figure 3 explains the two situations clearly. The true negatives by test 1 are added with true negatives by test 2, among test positives of test 1, there is gain in net specificity as a consequence of sequential screening test.

Example

Consider a hypothetical population of 5000 with prevalence rate of a specific disease as 10%. A preliminary screening test was applied on them with 70% sensitivity and 80% specificity. Those who tested positive were further tested by more powerful test with sensitivity 80% and specificity 90%. Estimate the net sensitivity and net specificity after two stage sequential screening tests.

To solve the above problem in sequential screening test following the above mentioned notations we have: n(S)= 5000; prevalence=10%; $n(S_1)$=(10÷100)×5000=500= those who have the disease; $n(S_2)$=5000– 500=4500= those who do not have the disease. Based on sensitivity and specificity of test1 we draw a 2×2 contingency table (Table 1).

Table 1: Result of test 1 and the disease status in sequential screening test.

Result of test 1	Disease		Total
	Positive	Negative	
Positive	350	900	1250
Negative	150	3600	3750
Total	500	4500	5000

Table 2: Result of test 2 among the positives of test 1.

Result of test 2	Disease		Total
	Positive	Negative	
Positive	280	90	370
Negative	70	810	880
Total	350	900	1250

$n(A \cap S_1)=(70 \div 100)\times 500=350$=true positive by test 1. $n(A^c \cap S_2)=(80 \div 100)\times 4500=3600$=true negative by test 1. $n(A \cap S_2)=4500-3600=900$=false positive by test 1. Thus n(A)=test positive by test 1=350+900=1250. These 1250 test positives were further tested by test 2 with sensitivity 80% and specificity 90% (Table 2).

$n(B \cap A \cap S_1)$=true positive by test 2 among test positives by test 1= $(80 \div 100)\times 350=280$; $n(B^C \cap A \cap S_2)=(90 \div 100)\times 900=810$=true negative by test 2 among test positives by test 1. Thus among 5000 population with 500 as having the disease, 280 are correctly diagnosed as positive after two stage sequential screening tests. Thus net sensitivity $=(n(B \cap A \cap S_1) \div n(S_1))\times 100=(280 \div 500)\times 100=56\%$. There is loss of net sensitivity as a result of sequential screening test over the sensitivity of both the individual tests.

Test negatives by test 1 were not further tested and therefore the true negative is $n(A^C \cap S_2)=3600$. Further, among the test positives by test 1, 810 were found true negatives by test 2. Therefore net specificity=(True negative by test 1+True negative by test 2 among positives by test 1) ÷ (Those who do not have the disease)×100=$[n(A^C \cap S_2) + n(A \cap B^C \cap S_2)] \div n(S_2)$ = (3600+810) ÷ 4500×100=98%. Thus as a result of two-stage sequential screening test net specificity has increased over the specificity of two individual tests.

Simultaneous screening tests

In simultaneous screening tests the results of two tests are independent. To be considered a positive case a test must be positive by either of the tests or by both tests.

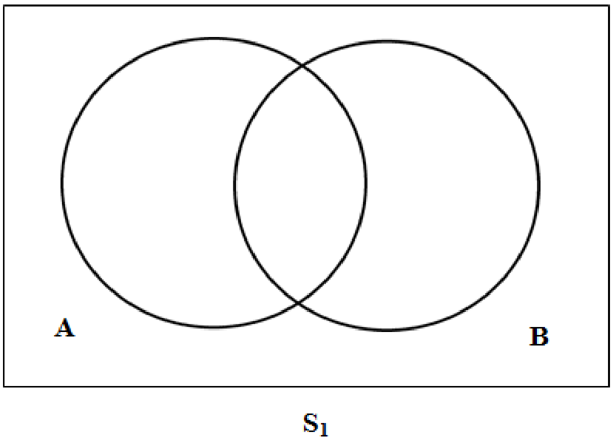


Figure 4: Venn diagram showing test positives by either of the tests in simultaneous testing.

We consider the following notations: S_1=set of those who have the disease. $n(S_1)$=total number of persons with the disease, based on the prevalence rate of the disease. A=set of positives by test 1; n(A) being their number. B=set of positives by test 2, n(B) being their number. $A \cap B$=set of positives by both tests. This is determined by applying the sensitivity of test 2 on test positives by test 1 and their number is denoted by $n(A \cap B)$.

Consider Figure 4. AUB=set of positives by either of the tests=A+B − $A \cap B$ (by union law of set theory). Hence, n (AUB)=n (A)+n (B) − n $(A \cap B)$. Thus net sensitivity=[n (AUB)) ÷ n (S_1)]×100.

For net specificity, a person is considered a negative case, if he is identified as negative by both the tests. We use the following notations: S_2=set of those who do not have the disease and $n(S_2)$ is their number. $n(S_2)=n(S) - n(S_1)$. C=set of negatives by test 1 based on specificity of test 1. n (C)=number of test negatives by test 1. D=set of negatives by test 2 based on specificity of test 2. n (D)=number of test negatives by test 2.

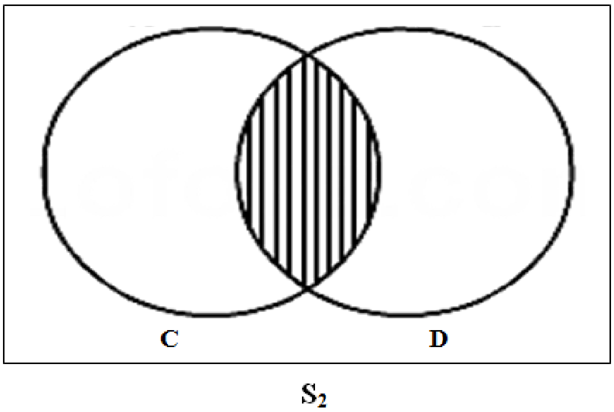


Figure 5: Venn diagram showing test negatives by both the tests in two simultaneous screening tests.

Consider Figure 5. $C \cap D$=set of test negatives by both tests determined by applying specificity of test 2 on the test negatives by test 1. The shaded area as shown in the Venn diagram and n $(C \cap D)$ is their number. Net specificity=(test negatives by both tests ÷ those who do not have the disease)×100=[n $(C \cap D)$] ÷ $n(S_2)$×100. There will be a loss in net specificity after simultaneous screening tests.

Example

In a population of 10000, the prevalence rate of a specific disease is 15%. Two simultaneous screening tests were conducted with sensitivity and specificity as given below. Sensitivity and specificity of test 1 and test 2 are 75% & 80% and 80% & 90% respectively. Estimate the net sensitivity and specificity after conducting two simultaneous screening tests.

To solve the above problem, we have: n (S)=10000; Prevalence of the disease = 15%; n (S_1)=those who have the disease=1500; n (S_2) = those who do not have the disease=8500; n (A) = positive by test 1 = (75÷100)×1500=1125; n (B) = positive by test 2 = (80÷100)×1500=1200; applying sensitivity of test 2 on test positives by test 1, one gets: n (A∩B) = (80÷100)×1125=900 = positive by both tests. Thus, number of positives either by test 1 or test 2 or both = n (AUB) = n (A)+n (B)─n (A∩B) = 1125+1200 ─ 900=1425. Hence net sensitivity = [n (AUB)]÷n (S_1)×100 = (1425÷1500) ×100=95%. Obviously there is gain in net sensitivity over the two tests after simultaneous screening tests.

Regarding estimation of net specificity, a person is considered negative, if he is tested negative by both the screening tests. We have, n (S_2) = those who do not have the disease=8500. Applying specificity of test 1, n (C) = those who test negative by test 1 = (80÷100)×8500= 6800. Applying specificity of test 2 on these 6800 samples one gets n (C∩D) = those who test negative by both tests = (90÷100)×6800=6120. Hence net specificity = [n (C∩D)÷n(S_2)]×100 = (6120÷8500)×100 = 72%. So there is loss in net specificity compared to those two tests after two simultaneous screening tests.

Funding: No funding sources
Conflict of interest: None declared
Ethical approval: Not required

REFERENCES

1. Gordis L. Epidemiology, 4th ed, Philadelphia: Saunders; 2008: 87-89.
2. Kanchanaraska S. Evaluation of diagnostic and screening tests, validity and reliability. John Hopkins Bloomberg. School of Public Health. John Hopkins University; 2008.

Cite this article as: Mukhopadhyay BB, Bhattacharjya H. Set theoretic approach to the concept of net sensitivity and net specificity in screening test results. Int J Community Med Public Health 2018;5:1690-3.