Sensitivity and Specificity

John Bullman

Emeritus Reader,

Eastman Dental Institute, London, UK

To anyone concerned with screening for medical or dental conditions, these terms are part of their stock in trade; but to the rest of us they may leave a feeling that we ought to know what they mean, but we are not absolutely sure. The purpose of this article is therefore to try to explain both their meaning and their importance in a way which a non-statistician can readily follow.

Screening is basically a procedure used to detect the presence of a disease or condition in groups of individuals; ideally long before the subjects concerned start to exhibit definite symptoms. In this way, treatment can be commenced earlier and so, it is hoped, be simpler, more effective or cheaper. Screening is designed to be applied to large groups rather than to individuals, so the screening methods employed need to be quick, easy to use, inexpensive and reliable.

Although, on the face of it, screening would seem to offer a useful and even life-saving answer to many medical and dental diagnostic problems, there are a number of both practical and ethical problems which need to be resolved before it can be employed. For example, to be cost-effective, the target condition either needs to have a high prevalence in the population, or, if the prevalence is low, it needs to be of high severity to the point of being life-threatening. On the ethical side, it would be wrong to institute screening if, for example:

- a) There were insufficient or inadequate resources available to treat all cases revealed by the process; or
- b) no adequate cure for the condition was available; or
- c) treatment existed, but did not necessarily confer benefit to the patient.

This can lead to difficult decision-making, when situations a, b or c exist, but there is nevertheless a desire to identify potential disease carriers in order to protect other members of the community.

However, the most relevant point to be considered here is the reliability of a screening test. An ideal test for any disease or condition will correctly identify each and every person who has the disease or condition, and will ignore those who do not have it. Sadly, this is an ideal which is rarely, if ever, achieved. Some tests are able to successfully identify every affected person, but only by also returning a false positive diagnosis on some who are not affected. In a perhaps worse scenario, other tests produce

negative results on all who are free from the condition, but also fail to identify all who are affected. There are thus two yardsticks which may be used to measure the efficacy of a screening test:

- To what extent is it successful in identifying all subjects with the target disease or condition?
- To what extent is it successful in identifying all subjects who are free from the target disease or condition?

At this point we can now introduce the two terms which form the heading for this article. For any given screening test:

- Sensitivity is the probability that a subject with the disease will screen positive.
- Specificity is the probability that a subject who is disease-free will screen negative.

How this works in practice is perhaps best illustrated by an example. Suppose we wish to screen 2000 individuals, of whom 1500 have the disease and 500 do not. Suppose further that a screening test correctly identifies 1400 of those with the disease, and also correctly identifies all but 50 of those who are unaffected. It is now possible to draw up a 2×2 contingency table to illustrate these results as in Table 1.

In this table the 'Disease Positive' column shows the breakdown of the 1500 subjects who were subsequently found to have the disease, while the 'Disease Negative' column shows the breakdown of the 500 subjects who were subsequently found to be disease-free. Similarly, the 'Test Positive' row shows the breakdown of the 1450 subjects who tested positive, and the 'Test Negative' row shows the breakdown of the 550 subjects who tested negative.

Table 1 Disease Screening

	Disease Positive	Disease Negative
Test Positive	1400	50
Test Negative	100	450
Total	1500	500

In this example, the sensitivity of the test will be $1400/1500 = 0.93$ or 93%, and the specificity of the test will be $450/500 = 0.90$ or 90%.

This means that there is a 93% probability that someone with the disease will screen positive, and a 90% chance that someone who is disease-free will screen negative.

A common misapprehension is to assume that just because a test has a sensitivity of 93%, this means that someone

who screens positive has a 93% chance of having the condition. To understand why this is not the case, consider the contingency results in Table 2 which is based on a test with a sensitivity of 95% and a specificity of 70%.

Sensitivity = $19/20 = 0.95$ or 95%. However, from the table it can be seen that the total who screened positive was $19 + 24 = 43$. Of these, only 19 actually had the disease. Thus if a person screened positive, the probability that

they had the disease was only $19/43 = 0.44$ or 44%; which is very different from the sensitivity of 95%.

In fact, the probability of a subject having the disease if they screen positive depends not only on the sensitivity but also on both the specificity and the prevalence of the disease in the total population.

Two other terms may be mentioned here. The positive predictive value is the probability that an individual who screens positive actually has the disease, and the negative predictive value is the probability that an individual who screens negative does not have the disease. In the worked example in Table 3, the positive predictive value is 0.44 or 44%.

$$\text{Sensitivity} = a/(a+c)$$

$$\text{Specificity} = d/(b+d)$$

$$\text{Positive predictive value} = a/(a+b)$$

$$\text{Negative predictive value} = d/(c+d)$$

$$\text{Prevalence} = (a+c)/n$$

It is useful to note that the concepts discussed above may also be employed with advantages when assessing the reliability of clinical examiners in diagnosing conditions under review in clinical trials or epidemiological investigations.

Table 2 Test with a sensitivity of 95% and specificity of 70%

	Disease Positive	Disease Negative
Test Positive	19	24
Test Negative	1	56
Total	20	80

Table 3 Summary

	Disease Positive	Disease Negative	Total
Test Positive	a	b	a+b
Test Negative	c	d	c+d
Total	a+c	b+d	n