Journal of Medical
Screening

Leader

Reporting of screening results

Simple and clear terminology is essential in the reporting of screening results. Historically, the reporting of screening results as positive and negative has been established for at least 40 years.[1] Recently, the use of this terminology has been questioned, mainly because of a concern that a positive result might be taken to mean having the disorder for which the test has been performed, and that a negative result might be taken to mean unaffected. This issue is particularly relevant to screening tests rather than diagnostic tests, because of the much reduced chance of being affected with a positive result.

Some tests are, by their nature, categorical, in that their findings can only be placed in categories, such as normal or abnormal, positive or negative; the results of other tests cannot be classified automatically in this way because they take a numerical value on a continuous scale. For example, a categorical test might be a mammogram in screening for breast cancer. An example of a test on a continuous scale is maternal serum α fetoprotein measurement in antenatal screening for open neural tube defects. In tests based on continuous variables there is rarely a natural level that separates test results into those that are positive and require further action and those that are negative and do not.

Positive and negative are standard terminology, widely used in practice and in most texts on screening.[2–7] They have the advantage of simplicity, clearly dichotomising the screened population into mutually exclusive categories, which is the intention. Much screening methodology relies on the use of this language—for example, monitoring the screen positive rate and the false positive rate. The use of the terms *screen* positive and *screen* negative helps to emphasise that the test is a screening test and the result should not be interpreted as diagnostic. The concern that the terms positive and negative may imply the presence or absence of the disorder being screened for needs to be allayed by prior explanation. This should include the use of well prepared information leaflets, pointing out, for example in Down's syndrome screening, that even though a woman may have a screen positive result, the probability that she has an affected pregnancy is, on average, only about 2%. Likewise, those undergoing screening need to know that a screen negative result carries a residual risk because screening does not detect all affected individuals.

The use of the terms high risk and low risk may have the advantage of appearing less categorical, suggesting that individuals in the high risk group may not have the disorder being screened for and that there is a residual risk in the low risk group, but there are disadvantages with this terminology. The terms high and low risk beg the question as to the definition of high and low. The same problem arises with higher risk and lower risk, which have sometimes been suggested as alternatives. A risk of 2% may not be regarded as high, yet this is the average risk in screen positives in many antenatal Down's syndrome screening programmes. A woman classified as having a "high risk" result may have a risk that is not particularly high—for example, a risk of 0.5% for Down's syndrome, and the term could cause needless worry. Similarly, a woman with a "low risk" may have a risk that is twice that of women in general.

Other possibilities for naming the group for which further action is being recommended include "recommend for recall" or "follow up rate". The former would be confusing because the expression is used in some screening programmes to describe those individuals being recalled for a subsequent round of screening—for example, in breast cancer screening. The latter can be confused with monitoring the programme and determining outcomes. The use of "further action offered" or "no further action offered" overcomes some of the difficulties, but the phrases do not lend themselves to quantification in the way that, for example, "false positive rate" does. The "further action offered" rate would be unwieldy, and relates to all positives rather than false positives only.

A practical advantage of the terms positive and negative is that they can be prefixed by "true" and "false" when relating the results of the test to the clinical outcome. In this way, terms such as "false positive" and "false negative" are accepted and understood. If the terms positive and negative were abandoned much of the prior literature on screening would lose its meaning.

It has been suggested that the problem can be avoided by not dividing the screenees into groups at all. Everyone who is screened is simply given a risk estimate, without categorisation. The main objection to this is that it would make the screening process relatively unpredictable, and a key element in the specification of a screening programme is to set a general policy in which the offer of a procedure or treatment is limited to those at greatest risk on the grounds of hazard or cost. It would be difficult to predict and monitor the efficacy and safety of a programme in which results are not categorised because the uptake of diagnostic testing is not known. Also, people probably find it helpful to know what judgment has been made collectively about where to draw the line. There are hazards in screening and it is widely felt that further action should be offered to those at greatest risk of the disorder for which they are being screened.

The problems of understanding screening results are more likely to arise from lack of information about the screening test than from the terminology used. It is more appropriate to improve the quality of information provided before screening than to alter useful terminology.

On balance, the terms screen positive and screen negative are better than the alternatives. The possibly false impression that a positive result necessarily means that the disorder being screened for is present and a negative result

means it is absent is best rectified by the education of health professionals, who can then provide clear information to individuals considering a screening test.

NICHOLAS J WALD

Department of Environmental and Preventive Medicine,
Wolfson Institute of Preventive Medicine,
St Bartholomew's and the Royal London School of Medicine and Dentistry,
Charterhouse Square, London EC1M 6BQ, UK

ELIZABETH R DORMANDY

Psychology and Genetics Research Group,
The Guy's King's and St Thomas's School of Medicine,
5th Floor Thomas Guy House,
Guy's Campus, London SE1 9RT, UK

1 Commission on Chronic Illness. *Chronic illness in the United States*. Vol 1. *Prevention of chronic illness*. Cambridge, Mass: Harvard University Press, 1957.
2 US Department of Health, Education, and Welfare. *Principles and procedures in the evaluation of screening for disease*. Washington DC: US Government Printing Office, 1961. (Public Health Monograph No 67. Public Health Services Publication No 846.)
3 Wilson JMG, Jungner G. *Principles and practice of screening for disease*. Geneva: World Health Organisation, 1968.
4 McKeown T. Validation of screening procedures. In: *Screening in medical care: reviewing the evidence*. Oxford: Oxford University Press for the Nuffield Provincial Hospitals Trust, 1968.
5 Whitby LG. Screening for disease: definitions and criteria. *Lancet* 1974;ii:819–21.
6 Holland WW, Stewart S. *Screening in health care*. London: Nuffield Provincial Hospitals Trust, 1990.
7 Miller AB, Chamberlain J, Day NE, *et al*. Report on a workshop of the UICC project on evaluation of screening for cancer. *Int J Cancer* 1990;**46**: 761–9.