Risk stratification in medical screening

Editorial

Nicholas Wald, University College London

Stephen Duffy, Queen Mary University of London

Allan Hackshaw, University College London

It is standard practice in medical screening to set a screening cut-off level that dichotomises results into screen-positive and screen-negative. This cut-off could be applied to the measurement of a single marker, such as alpha-fetoprotein in antenatal screening for neural tube defects,[1] or to the risk calculated using a combination of multiple markers, for example in antenatal screening for Down's syndrome[2] or QRISK® in cardiovascular disease screening.[3] Screening may be justified when the subsequent diagnostic investigations of interventions are harmful, costly, or both, and there is good discrimination between individuals who are (or will be) affected by a disorder and those who are not. This discrimination is often referred to as screening performance and can be expressed in terms of the detection rate (sensitivity) for a given false-positive rate. If the intervention is effective for all, safe and inexpensive, primary prevention is the strategy of choice instead of screening.

In recent years, there has been growing interest in what is referred to as 'risk stratification' in screening. The aim is to improve screening performance. There has, however, been a lack of clarity about what is meant by risk stratification in medical screening and whether it improves screening performance. Some have, incorrectly, used the term risk stratification to describe the use of additional screening factors such as polygenic risk scores and family history[4] to determine eligibility for, say, mammography in breast cancer screening, but this process is itself a screening enquiry which forms part of step-wise screening. Low-dose pulmonary CT scanning offered on the basis of smoking history is another example of step-wise screening.[5] To help clarify the meaning of risk stratification in the context of medical screening we suggest the following definition:

> The categorisation of individuals who are screened for a disorder into strata of risk based on an initial screening test or enquiry with the intention of offering different screening protocols or intervention policies according to the stratum in which an individual is placed.

An example of risk stratification is the MyPeBS study[6] of breast cancer screening, which uses risk stratification to allocate women into four groups to offer different subsequent screening protocols rather than different interventions:

- Low risk: four-yearly mammography;
- Moderate risk: two-yearly mammography;
- High risk: annual mammography;
- Very high risk: annual mammography and MRI.

As defined, and illustrated by this example, neither multiple-marker screening which uses risk as the screening variable or step-wise screening is an example of risk stratification. If the

intervention among people with positive screening results is costly or harmful, risk stratification might be considered provided the range of risk across the strata is large.

But there are disadvantages with risk stratification. If there are only a few strata there may be many people in the highest risk stratum, most of whom will be false-positive. Also, using the breast cancer screening example, women in the 'very high risk' category and therefore offered more frequent mammograms could reasonably interpret their 'result' as being positive and as a consequence be acutely worried. Similarly, those in the lowest risk category may be falsely reassured. Risk stratification therefore has the potential to lead to an increase in the anxiety associated with being a false-positive but with little gain in detection. The issue is a topic that is being investigated.[5]

Another disadvantage is the complexity of risk stratification and its perception by potential screenees. Individuals would need to understand the stratification of results; if one stratum is high risk and another intermediate risk with different consequences, it could cause annoyance and distress among people in the intermediate stratum who are found to have the disorder being screened for. Also, if the difference in risk between the different categories is relatively small, the simplicity of screening based upon having a simple positive or negative result should probably not be sacrificed by adding to the screening process a precision which is of little clinical significance. There would need to be discussions over what this difference should be; it would probably differ for different screening programmes depending on cost and potential harms.

It is recognised that 'early detection' is not always worthwhile in spite of its intuitive appeal; the same applies to risk stratification. A risk stratification policy that is only less expensive to deliver (i.e., more cost-effective) may be difficult to justify clinically if many affected individuals are not detected. Any application of risk stratification in medical screening would have to show a clinically significant improvement in screening performance to justify its implementation.

References

1. Wald NJ, Cuckle H, Brock JH, et al. Maternal serum-alpha-fetoprotein measurement in antenatal screening for anencephaly and spina bifida in early pregnancy. Report of U.K. collaborative study on alpha-fetoprotein in relation to neural-tube defects. *Lancet* 1977; 1: 1323–1332.

2. Wald NJ, Hackshaw AK, Walters J, et al. First and second trimester antenatal screening for Down's syndrome: the results of the serum, urine and ultrasound screening study (SURUSS). *J Med Screen* 2003; 10: 56–104.

3. Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. *Br Med J* 2017; 357: j2099.

4. Garcia-Closas M, Gunsoy NB, Chatterjee N. Combined associations of genetic and environmental risk factors: implications for prevention of breast cancer. *J Natl Cancer Inst* 2014; 106: dju305.

5. Lung Cancer Screening Programme: Cancer Research UK, https://www.cancerresearchuk.org/health-professional/cancer-screening/lung-cancer-screening (2024, accessed 26 April 2024).

6. Roux A, Cholerton R, Sicsic J, et al. Study protocol comparing the ethical, psychological and socio-economic impact of personalised breast cancer screening to that of standard screening in the "my personal breast screening" (MyPeBS) randomised clinical trial. *BMC Cancer* 2022; 22: 507.