

Screening for serious illness

Limits to the power of medicine

Juan Gérvas

Health and disease are conceptualised in both the medical and popular culture as value-free objective constructs. They are taken to represent real states that can be objectively determined. However, they are also portrayed in ways that incorporate evaluative judgements such that a diseased state is seen as abnormal, dysfunctional and bad.¹ Furthermore, there is an evaluative system derived from science which holds that what derives from patient experience is 'subjective' and inherently untrustworthy whereas that deriving from diagnostic procedures and the observations of doctors is 'objective' and, therefore, an infallible representation of 'reality'. The consequences of these oversimplifications are nowhere more apparent than in the case of screening for serious illnesses where belief in the 'reality' of disease derived from screening procedures is assumed to always outweigh the patient's experience of apparent health. This can be illustrated in respect of two important malignant diseases for which screening is widely advocated - melanoma and breast cancer.

Screening for melanoma - not such a benign practice?

No randomised trials have yet been reported to show any survival benefit from early detection of melanoma. However, there are advocates for active surveillance of the population for melanoma. Increased surveillance for melanoma has led to a 'melanoma anxiety dynamic' with more and more biopsies on progressively less suspicious lesions lowering the positive predictive value of both formal clinical and histological evaluation.² Contrary to popular belief, the phenotypic features of pigmented lesions are not dichotomous traits but rather continuous variables along a scale. Histopathology is, by nature, qualitative and subjective.^{3,4} Thus, one study of Italian pathologists found a κ value of 0.61 in a comparison of the performance of pathologists in distinguishing cutaneous melanoma from benign lesions.⁵ A total of 140 slides (120 originally diagnosed as melanoma and 20 as a benign lesion) were

circulated to four histopathologists with at least ten years experience as dermatopathologists. There was concordance between all pathologists on the diagnosis of 85 cases of malignant melanoma (70.8%). The overall agreement suggests considerable variability in the histopathological diagnosis made by single pathologists. One expert thought that there were 119 melanomas (99.1%) whereas another expert considered 88 as melanomas (73.3%).⁵ Morphological diagnosis is entirely subjective and errors in diagnosis are inevitable; physicians know it and patients should be informed.⁶ In a similar US study, eight acknowledged expert dermatopathologists each submitted five cases of melanoma or melanocytic nevi they considered to be classic cases. Each pathologist was then sent 37 slides (including two extra added by the researchers). The experts agreed unanimously about only 11 of 37 melanocytic neoplasm (30%). In 62% of cases there was unanimous agreement or only one discordant designation. The κ value was 0.50.⁷ Application of histological criteria is a subjective process: one expert thought that there were 21 malignant neoplasms and 16 benign nevi, whereas another expert considered 10 to be malignant, 26 benign, and one indeterminate. The US Preventive Services Task Force noted in 2001, regarding skin cancer, that 'there is still insufficient scientific evidence to determine whether regular total body skin examination for skin cancer is effective in reducing illness and death', the same conclusion the Task Force had reached in 1996.^{8,9}

Breast cancer screening - instilling a protective awareness or breeding unjustified fear?

Most breast cancers occur in older women but in raising awareness, cancer organisations, the media and even medical journals tend to portray case histories involving young women.¹⁰ A life table analysis of the cumulative incidence of breast cancer in England and Wales shows that the risk for women under age 35 is 1 in 625.¹¹ It is true that women have a 1 in 12 risk (in the UK) or a 1 in 8 risk (in the US) of developing breast cancer but only if they manage to escape other threats to life and survive to the age of 80. Incidence is not equivalent to mortality: In England and Wales, only one woman in 26 will have died of breast cancer by the age of 80.¹¹ In the US, breast cancer mortality has remained constant since 1940 while incidence has

Juan Gérvas, general practitioner.

Canencia de la Sierra, Equipo CESCA, Madrid, Spain.

Visiting professor, Johns Hopkins School of Public Health, Baltimore, US.

Submitted: 26 June 2001; accepted (in revised form): 29 May 2002.

nearly doubled.¹² The increased incidence may reflect overdiagnosis of tumours of low grade malignancy or of carcinoma *in situ*.¹¹ Age-adjusted annual breast cancer death rates in 1990-1993 vary from 6.6 per 100,000 in Japan to 27.7 in the UK (15.1 in Greece, 16.6 in Finland, 17.3 in Spain, 17.7 in Sweden, 22 in the US, 26.8 in Ireland and 26.9 in the Netherlands).¹³

Mammography is the lynchpin of screening for breast cancer, yet the question of whether detection by mammography contributes to the life expectancy in those in whom breast cancer is confirmed remains moot.¹⁴ Radiologists can differ, sometimes substantially, in their interpretation of mammograms and in their recommendations for management. In the US, ten radiologists who routinely read mammograms and who were unaware of the diagnosis and research design were invited to participate in a study about variability in radiologists' interpretations of mammograms.¹⁵ In phase 1 of the research, each of the ten radiologists independently read the mammograms of 150 patients (27 from women with histopathological confirmed breast cancer and 123 from women with no evidence of breast cancer after three years of follow-up examinations). In phase 2, which occurred five months later, the same films were reviewed again, this time in a new, randomly arranged sequence (the radiologists were not told that they would be seeing the same films a second time). All ten radiologists agreed on the diagnostic interpretation in only ten cases (7%). The median agreement was 78% (κ value 0.47) for inter-observer variability in diagnostic interpretation and 85% (κ value 0.49) for the recommendation of biopsy (the corresponding numbers for intra-observer variability were 84%, 0.57 and 91%, 0.71). The ten radiologists differed widely in their recommendations for management; for example, radiologist A recommended an immediate work-up for almost all the patients with cancer (96%), but also for 64% of the patients without cancer. When two or more radiologists recommended a biopsy of the same patient, a disagreement in the stated location (right or left breast) occurred in 2% of the paired comparisons among the radiologists but in 9% of comparisons of women as a whole. The choice of a screening threshold for identifying potential malignancies may reflect personal estimates of the probability of cancer but also subjective factors.¹⁵

About 5% of screening mammograms are positive or suspicious, and of these 80-93% are false positives that cause much unnecessary anxiety and further procedures including surgery.^{14,16} The problem is compounded by the inclusion of cases of carcinoma *in situ*, whether ductal or lobular, natural history of which is not well understood.^{17,18} In the Nijmegen Study, the definition of cancer varied and lobular carcinoma *in situ* was not counted as cancer; i.e., a mammogram resulting in the finding of lobular carcinoma *in situ* was considered to

be a false-positive screen (other studies have included carcinoma *in situ* as 'real' cancer).²⁶ The cumulative risk of a false-positive result of breast cancer screening test was 49.1% after ten mammograms in an American study; the authors estimated that among women who do not have breast cancer, 18.6% will undergo a biopsy after ten mammograms.¹⁹

The false-negative reassurance of a negative mammogram is another serious issue since 10-15% of early breast cancers are missed by mammography: 'there is no justification for the argument that one of the advantages of screening mammography is the comfort and reassurance that cancer is not present'.¹⁴ The screening threshold for identifying potential malignancies will influence not only the proportion of cancers detected but also the likelihood that invasive diagnostic procedures will be deemed necessary and productive.¹⁶

Conclusion

The scientific base of medicine is weak and it will be better for everybody if that fact were more widely recognised. We need to understand the extent of our ignorance and share it with the public, patients and policymakers.²⁰ No diagnosis/treatment process is without its dangers, and if doctors use it where they have no solid scientific evidence of benefit they are exposing patients to risk when there may be no benefit. People invited for screening of serious diseases, as melanoma and breast cancer, must be told about the risks, benefits and limitations in a way that instils realistic expectations and ensures fully informed consent in those who participated.²¹ General practitioners have a duty, which goes beyond the law, to make sure that they have the patient's permission for whatever they are proposing to do. The relationship between the medical profession and the public is changing; 'what has not changed is the fact that the public need doctors who are knowledgeable and skilled, ethical and committed'.²² ■

This editorial is based on the lecture 'Early diagnosis of serious diseases: screening, health and medical power' given by Dr Gervas at the 2001 WONCA Europe Congress in Tampere.

References

- 1 Chiong W. Diagnosing and defining disease. *JAMA* 2001;285:89-90.
- 2 Esch EPV, Muir CS, Nectoux J, MacFarlane G, Maisonneuve P, Bharucha H, et al. Temporal change in diagnostic criteria as a cause of the increase in malignant melanoma is unlikely. *Int J Cancer* 1991;47:483-9.
- 3 Edman RL, Klaus SN. Is routine screening for melanoma a benign practice? *JAMA* 2000;284:883-5.
- 4 Piepkorn M, Odland PB. Quality of care in the diagnosis of melanoma and related melanocytic lesions. *Arch Dermatol* 1997;113:1393-6.
- 5 Corona R, Mele A, Amini M, Rosa G, Coppola G, Piccardi P, et al. Interobserver variability on the histopathologic diagnosis of cutaneous melanoma and other pigmented skin lesions. *J Clin Oncol* 1996;14:1218-23.
- 6 Ackerman AB. Discordance among expert pathologists in diagnosis of melanocytic neoplasms. *Hum Pathol* 1996;27:1115-6.

- 7 Farmer ER, Gonin R, Hanna MP. Discordance in the histopathologic diagnosis of melanoma and melanocytic nevi between expert pathologists. *Hum Pathol* 1996;27:528-31.
- 8 U.S. Preventive Service Task Force. Guide to clinical preventive services. 2nd ed. Baltimore: Williams & Wilkins, 1996.
- 9 <http://www.ahrq.gov/clinic/prevenix.htm>.
- 10 Kent A. Raising awareness or spreading fear? *BMJ* 2000;321:247.
- 11 Bunker JP, Houghton J, Baum M. Putting the risk of breast cancer in perspective. *BMJ* 1998;317:1307-9.
- 12 Harris JR, Lippman ME, Veronesi U, Willet W. Breast cancer. *N Engl J Med* 1992;327:319-28.
- 13 Shapiro S, Coleman EA, Broeders M, Codd M, Koning H, Fracheboud J, et al. Breast cancer screening programmes in 22 countries: current policies, administration and guidelines. *Int J Epidemiol* 1998;27:735-42.
- 14 Wright CJ, Mueller CB. Screening mammography and public health policy: the need for perspective. *Lancet* 1995;346:29-32.
- 15 Elmore JG, Wells CK, Lee CH, Howard DH, Feinstein AR. Variability in radiologists' interpretations of mammograms. *N Engl J Med* 1994;331:1493-9.
- 16 Mushlin AI, Kouides RW, Shapiro DE. Estimating the accuracy of screening mammography: a meta-analysis. *Am J Prev Med* 1998;14:143-53.
- 17 Silverstein MJ, Lagios M. Benefits of irradiation for DCIS: a Pyrrhic victory. *Lancet* 2000;355:510-1.
- 18 Ernster V, Barclay J, Kerlikowske K, Wilkie H, Ballard-Barbash R. Mortality among women with ductal carcinoma in situ of the breast in the population-based surveillance, epidemiology and end results program. *Arch Intern Med* 2000;160:953-8.
- 19 Elmore JG, Barton MB, Moceri VM, Polk S, Arena PJ, Fletcher SW. Ten-year risk of false positive screening mammograms and clinical breast examinations. *N Engl J Med* 1998;338:1089-96.
- 20 Smith S. The ethics of ignorance. *J Med Ethics* 1992;18:117-8,134.
- 21 Wilson RM. Screening for breast and cervical cancer as a common cause for litigation. *BMJ* 2000;320:1352-3.
- 22 Irvine D. The changing relationship between the public and the medical profession. *J R Soc Med* 2001;94:162-9.

Continued from page 46.

which are associated with an illness are present we start treatment and send for tests to confirm our impression. Gulich's algorithm is a more formalised version of this approach which has addressed carefully the best combination of signs to ensure that the C-reactive protein test is used to the best advantage.

A more accurate method to use is to carry out a Bayesian analysis, multiplying the likelihood ratio for presence or absence of each data item by the prior probability to produce a posterior probability. A simpler way to do this in practice is to use a Bayesian-scoring system such as the B-score system,² where a score for presence or absence of each data item is added together to produce a total score which can be converted into the probability of the diagnosis being present. Test results can be easily incorporated into a system such as this when the likelihood ratios for positive and negative test results are known. A decision on whether to use another test can be made easily, based on whether the score resulting from a positive or negative test result could change the overall probability sufficiently to change the management of the patient. Cheap tests with high scores can be used first to achieve cost-effectiveness. Different prior probabilities due to varying incidences in different populations, and selection by referral to specialist clinics can also be allowed for.

So what?

The 'So what' question is very important in research. How accurately do we need to make a diagnosis? Self-limiting illness is very common in primary care. Often our role is to manage a condition in a symptom-treating mode until the body's natural repair mechanisms restore health. In the case of strep throat, management without antibiotics may even result in better development of resistance and less risk of recurrence. But the rare life-threatening illness is always in the back of our minds. We have to be able to diagnose illnesses reliably which have a serious outcome if not treated with a specific curative method. We have to be able to refer when a condition is outside our expertise, and we have to be able to support our patients through passing on our knowledge and understanding, and providing a relationship which aids their emotional and behavioural coping mechanisms. ■

References

- 1 Gulich M, Triebel T, Zeitler H-P. Development and validation of a simple, two-step algorithm to identify streptococcal infection in adults with sore throat. *Eur J Gen Pract* 2002;8:57-61.
- 2 Dobbs F. A scoring system for predicting Group A streptococcal throat infection. *Br J Gen Pract* 1996;46:461-4.