Journal of Medical
Screening

Editorials

Populist instead of professional

The evidence on the efficacy of mammographic breast cancer screening is as firm as any can be in medical science. Randomised trials show clearly that breast cancer mortality is about a third lower in screened women. Many expert groups have reviewed the evidence and concluded that breast cancer screening is worthwhile. In spite of this a recent paper by Gøtzeche and Olsen[1] re-examined the same evidence and concluded that breast cancer screening is not justified. The paper was published in *The Lancet*, and the sensation of the claim, seemingly supported by so prestigious a journal, attracted considerable publicity in the media. Experts interviewed at the time reinforced the opinion that women should be screened but did not give clear reasons why the report was judged to be flawed, and so people remained confused.

Medical journals and the media, who take their cue from papers published in the medical journals, sometimes tend to encourage controversy rather than scholarship and sound interpretation. Sensational headlines and increased circulation become more important than accuracy and good judgment. Gøtzesche and Olsen's paper lacks scientific merit. The accompanying editorial by de Koning[2] stated the problems in the paper and one might ask why his comments were not referred back to the authors with a view to correcting it instead of publishing both in the same issue together.

The flaws in the paper published in *The Lancet* were covered in detail in the subsequently published correspondence.[3–9] Gøtzesche and Olsen considered seven randomised trials of breast cancer screening and inappropriately rejected five valid trials because they failed to acknowledge that minor chance differences in age and social class between screened and control groups will arise with the cluster randomisation used in most of these trials. Such differences, being random, will tend to cancel out when trials are combined. Their analysis rested on only two trials, one of which did not compare screening with no screening (it was a comparison of mammography plus breast palpation with breast palpation alone). The other trial was valid, but preliminary results were used rather than more recent results, and data on women under 50 and over 50 were combined: they should have been analysed separately because it remains uncertain whether mammography significantly reduces mortality in women under 50. To what extent the paper by Gøtzeche and Olsen has undermined the confidence of women invited to attend breast cancer screening remains to be seen. More generally, however, reports like this will tend to make the public sceptical of the quality and reliability of all medical findings reported by the media—to regard medical reports as no more than the "opinions" of one group that are probably going to be countered by the "opinions" of others. Medical science becomes medical opinion and the disagreement becomes entertainment.

The Lancet should not have published this paper. Editors of medical journals have a duty to try to publish valid work and to correct known errors of fact or interpretation before papers are published. This is particularly the case where the issue is an important public health measure that will affect millions of people. To publish a paper which the accompanying editorial rightly criticises as being unsound is being more populist than professional.

NICHOLAS WALD

Editor

1 Gøtzeche PC, Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet* 2000;**355**:129–34.
2 de Koning HJ. Assessment of nationwide cancer-screening programmes. *Lancet* 2000;**355**:80–1.
3 Duffy SW, Tabar L. Screening mammography re-evaluated. *Lancet* 2000;**355**:747–8
4 Moss S, Blanks R, Quinn MJ. Screening mammography re-evaluated. *Lancet* 2000;**355**:748.
5 Nyström L. Screening mammography re-evaluated. *Lancet* 2000;**355**:749.
6 Hayes C, Fitzpatrick, Daly L, *et al.* Screening mammography re-evaluated. *Lancet* 2000;**355**:749.
7 Law ML, Hackshaw AK, Wald NJ. Screening mammography re-evaluated. *Lancet* 2000;**355**:749–50.
8 Cates C, Senn S. Screening mammography re-evaluated. *Lancet* 2000;**355**:750.
9 Rozenberg S, Liebens F, Ham H. Screening mammography re-evaluated. *Lancet* 2000;**355**:751–75.

Screening for hypothyroidism in adults: supporting data from two population studies

Data from two observational population studies indicate that the time may be at hand to consider new strategies for identifying and managing hypothyroidism.

In the first of these studies, begun in 1972,[1–3] 2779 adult residents of Whickham, England, provided baseline health information and blood and urine samples. These study subjects were then followed up at intervals for 20 years. The primary purpose was to learn about the prevalence of the whole range of autoimmune thyroid disorders in communities, as opposed to selected hospital populations.

The second study in Maine, USA, focused on a cohort of 25 216 women who were pregnant between 1987 and 1990.[4 5] Serum samples were collected from these women during the second trimester and stored in the freezer. Thyrotrophin measurements were then performed on all those serum samples in 1996 and 1997, and 62 of the

women with raised thyrotrophin levels were selected for individual follow up, together with 124 matched controls. Thyroid peroxidase antibodies and free thyroxine measurements were also tested in this group of 186 women, and neuropsychological assessment was carried out on the children of these women at 7 to 9 years of age. The major purpose of this second study was to learn whether untreated or inadequately treated maternal hypothyroidism during pregnancy might adversely affect fetal brain development. In both groups, autoimmune thyroiditis was the major cause of hypothyroidism.

At the end of 20 years, investigators in the Whickham study calculated that 3.5 new cases of spontaneous hypothyroidism occurred per 1000 surviving women per year (95% confidence interval 2.8 to 4.5). Hypothyroidism in the study group was about six times more common in women than in men, and the rate of occurrence of new cases increased steadily with age. In Maine, five new cases of hypothyroidism were recorded among the 120 available control women after 10 years, yielding an estimate of 4.2 new cases per 1000 women per year. These two estimates are remarkably similar. Of even greater interest was the Whickham study finding that 55% of the women with raised thyrotrophin measurements and positive thyroid antibodies at baseline were clinically hypothyroid 20 years later. In Maine, 64% of the women with raised thyrotrophin measurements at baseline were hypothyroid 10 years later, and an average of five years elapsed between those baseline measurements and the time when a clinical diagnosis was made.

The high correlation between raised thyrotrophin levels, positive thyroid antibodies, and subsequent development of clinically apparent hypothyroidism recorded in the two studies leads to the conclusion that it may be feasible to use serum thyrotrophin measurements to identify hypothyroid subjects systematically at a presymptomatic or early symptomatic stage and pre-emptively treat or follow them up, thereby avoiding the clinical consequences. The prospect of such an approach is made more attractive by the reliability and low cost of thyrotrophin testing, and also by the safety and low cost of thyroid replacement. In 1981 the authors of the Whickham study explained the shortcomings of the traditional method of diagnosing hypothyroidism by the clinical evaluation of a doctor, as follows: "Overt hypothyroidism is a common condition that develops insidiously and is often not recognised until it has been present for a considerable time. The clinical features are not specific, and the diagnosis is often made only when there is a full range of the symptoms and signs that are associated with myxoedema."[2] The accuracy of this statement is borne out by the long delay in making a clinical diagnosis of hypothyroidism that was reported in the Maine study.

Insights gained from the Maine study argue in favour of providing thyrotrophin measurements to all pregnant women, regardless of age. The lengthy delay in making a clinical diagnosis in so many of the women with initial biochemical evidence of hypothyroidism means that their ability to function in everyday life is likely to be interfered with at a time when the demands placed on them during their child's early years are greatest. The fatigue and depression associated with hypothyroidism are particularly disabling in such circumstances. Early detection of thyroid deficiency through screening would allow treatment to be started before such clinical problems arose. The Maine study also identified a lower average IQ among the children of untreated hypothyroid mothers, and it would be an important additional advantage if early treatment were to avoid that problem, as well. Further studies are necessary to determine the efficacy of thyroid replacement in protecting that aspect of fetal development.

To be effective, any screening strategy needs to be thoroughly worked out, simple and straightforward, and consistently applied. Screening for hypothyroidism is no exception. Data from the Whickham study suggest that screening might be limited to women and that age might be used as an initial screening test in non-pregnant women. A thyrotrophin cut off point would need to be agreed for defining high risk, and decisions would need to be made as to when other thyroid testing, such as antibody measurements, ought to be used. Guidelines for treatment and follow up would need to be established and provision made for stopping treatment at some point to determine whether thyroid function had returned to normal. Time schedules for subsequent screening of initially euthyroid subjects might be based upon projected, age dependent rates of occurrence. Between them, the two studies provide abundant information to guide those responsible for making these types of policy decisions. All of this leads to the conclusion that pilot screening trials should be implemented in both pregnant and non-pregnant populations to determine the feasibility of this approach in everyday practice.

JAMES HADDOW

Foundation for Blood Research,
PO Box 190,
Scarborough,
Maine 04074,
USA

1 Tunbridge WMG, Evered DC, Hall R, *et al*. The spectrum of thyroid disease in a community: the Whickham Survey. *Clin Endocrinol* 1977;7:481–93.
2 Tunbridge WMG, Brewis M, French JM, *et al*. Natural history of autoimmune thyroiditis. *BMJ* 1981;**282**:258–62.
3 Vanderpump MPJ, Tunbridge WMG, French JM, *et al*. The incidence of thyroid disorders in the community: a twenty-year follow-up of the Whickham Survey. *Clin Endocrinol* 1995;**43**:55–68.
4 Klein RZ, Haddow JE, Faix JD, *et al.* Prevalence of thyroid deficiency in pregnant women. *Clin Endocrinol* 1991;**35**:41–6.
5 Haddow JE, Palomaki GE, Allan WC, *et al.* Maternal thyroid deficiency during pregnancy and subsequent psychological development in the child. *N Engl J Med* 1999;**341**:549–55.

Screening for type II diabetes mellitus

In the decade following the second world war screening for non-insulin dependent diabetes mellitus (type II according to current nomenclature) became popular in North America. An immediate problem was the lack of consensus about the criteria for diagnosing diabetes, in particular in asymptomatic subjects. Blood glucose levels, whether fasting, casual, or post-load, were approximately normal in distribution and there was no statistical or other evidence to separate diabetic from non-diabetic subjects. This problem was temporarily resolved by epidemiological studies that showed an apparent threshold of post-load glucose level above which the risk of diabetic microvascular disease

(principally diabetic retinopathy) became obvious. This was incorporated into recommendations made by the National Diabetes Data Group in the USA and by the WHO expert committee, refined by the WHO Study Group in 1985,[1] whose dicta were largely accepted.

In 1978 a consensus conference on screening for diabetes was convened by the Centers for Disease Control in Atlanta.[2] The conclusion was that community screening was not justified, the principal reason being that there was no evidence that treatment which lowered blood glucose diminished the risk of diabetic complications. This view was reinforced by the negative results of the first major clinical trial of hypoglycaemic treatment in type II diabetes mellitus carried out by the University Group Diabetes Programme. However, two more recent, and larger, trials have shown the undoubted benefits of hypoglycaemic treatment, particularly for microvascular disease in both types I and II diabetes mellitus.[3 4] These results have been reinforced by reports from smaller clinical trials. Also, various treatments have been shown to be as least as effective in diabetic patients as in those without diabetes on the extra cardiovascular risk associated with, if not due to, the diabetes.[5 6]

So, according to the Wilson and Junger criteria for screening, type II diabetes mellitus is a medically important disorder, which has a known (approximately) prevalence, for which there are effective treatments, and which can be ethically justified. The screening test is relatively simple and not expensive (though the details are either disputed or not adequately researched). These considerations have led the American Diabetes Association[7] to make some rather equivocal recommendations about screening. Thus, for example, screening "may be appropriate if the patient has one or more the risk factors shown in table 1". As one of those risk factors they list is age >45 years, the potential population is substantial. For community screening "there is insufficient evidence to conclude that community screening is a cost effective approach to reduce the morbidity and mortality associated with diabetes in presumably healthy subjects." Nevertheless "based on expert opinion, community screening for diabetes in a high risk population *may* (my italics) be worthwhile but its true efficacy is unknown." In the USA this seems to have been interpreted by doctors as providing carte blanche for general screening of those aged over 45.[8]

Any possible conclusions have been complicated by the further difference of opinion about the screening test. The American Diabetes Association experts have recently argued that fasting blood glucose values, rather than post-load values recommended by the WHO, should be used both for screening and diagnosis.[9] This has the advantage that subjects do not have to wait for two hours for the post-load sample, which tends to encourage doctors to screen more people. However, it appears that the fasting diagnostic values espoused by the American Diabetes Association are less sensitive than the post-load values recommended by the WHO.[10] A further complication, which affects the criteria of both the American Diabetes Association and the WHO, is that the extra risk of cardiovascular disease associated with hyperglycaemia becomes apparent at blood glucose levels below those used to define diabetes mellitus.[11] There is already debate as to whether blood glucose levels alone should be set in screening for cardiovascular disease rather than diabetes microvascular risk. A major argument in favour of this is that the *absolute* risk is greater for cardiovascular morbidity and mortality.

With regard to microvascular disease, observational studies and modelling exercises (both subject to caveats[12]) suggest that tight control of blood glucose in older subjects with moderate hyperglycaemia (on treatment) is of little benefit because of the low risk of significant microvascular complications. The principal argument for screening older subjects, therefore, would be to identify people at extra risk of cardiovascular disease. This presupposes, as does screening for the microvascular risk, that ordinary care can achieve the same results as those achieved in clinical trials, which is demonstrably not the case for treatment with insulin to lower blood glucose in type II diabetes mellitus.[12]

In conclusion, those looking for an evidence base to determine their attitude to screening for type II diabetes mellitus will have to wait for more and better evidence before reaching firm conclusions.

R J JARRETT

Emeritus Professor of Clinical Epidemiology,
University of London,
45 Bishopsthorpe Road,
London SE26 4PA, UK

1 WHO Study Group. *Diabetes mellitus.* Geneva: WHO, 1985. (Technical report series 727.)
2 Herron CA. Screening in diabetes mellitus: report of the Atlanta Workshop. *Diabetes Care* 1979;**2**:357–62.
3 The Diabetes Control and Complications Trial Research Group. The effects of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *N Engl J Med* 1993;**329**:977–86.
4 UK Prospective Diabetes Study (UKPDS) Group. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type II diabetes (UKPDS 33). *Lancet* 1998;**352**:837–53.
5 Hansson L, Zanchetti A, Carruthers SG, *et al*. Effects of intensive blood pressure lowering and low-dose aspirin in patients with hypertension: principal results of the Hypertension Optimal Treatment (HOT) randomised trial. *Lancet* 1998; **351**:1755–62.
6 Pyorala K, Pedersen TR, Kjekshus J, *et al*. Cholesterol lowering with simvastatin improves prognosis of diabetic patients with coronary heart disease. *Diabetes Care* 1997;**20**:614–20.
7 American Diabetes Association. Screening for type 2 diabetes. *Diabetes Care* 1999;**22**:S20–3.
8 Harwell TS, Smilie JG, McDowell JM, *et al*. Diabetes screening practices among individuals aged 45 years and over. *Diabetes Care* 2000;**23**:125–6.
9 American Diabetes Association. Report of the expert committee on the diagnosis and classification of diabetes mellitus. *Diabetes Care* 1999;**22**:S5–9.
10 Davies M. New diagnostic criteria for diabetes - are they doing what they should? *Lancet* 1999;**354**:610–11.
11 Shaw JE, Dowse GK. Zimmet PZ, *et al*. Impaired fasting glucose: how low should it go? *Diabetes Care* 2000;**23**:34–9.
12 Vijan S, Kent DM, Hayward RA. Are randomised controlled trials sufficient evidence to guide clinical practice in type II (non-insulin dependent) diabetes mellitus? *Diabetologia* 2000;**43**:125–30.

* * * * *

The challenge of preventive medicine in 2000

"We are under no illusion that preventive strategies will be easy to implement. For a start, the costs of prevention have to be paid in the present, while its benefits lie in the distant future. And the benefits are not tangible—when prevention succeeds, nothing happens. Taking such a political risk when there are few obvious rewards requires conviction and considerable vision."

Kofi Annan, Secretary General of the United Nations writing on "Preventing conflict in the next century", in the Economist publication, *The World in 2000*.[1]

K ANNAN

Secretary General of the United Nations

1 Annan K. Preventing conflict in the next century. In: Fishburn D, ed. *The world in 2000*. London: Economist Publications, 1999:91.