Periodic Health Examination –
A brief history and critical assessment

Walter Holland

Summary: The development of the periodic health examination (PHE) is outlined with a critical analysis of the evidence for its benefits to health improvement. The policies of governments in supporting the use of PHE are questioned. Evidence is presented that population policies are more effective in health improvement and reducing social-health inequalities. PHE is more an attractive populist measure and supported by commercial interests. The lack of effectiveness in improving health, increasing health inequalities and its unfortunate side-effects are ignored.

Keywords: periodic health examination, multiphasic screening, screening, history, public health policy

The start

D’Souza, stated that the origins of the Periodic Health Examination (PHE) or screening occurred in a brothel in the papal state of Avignon in 1347, “when an abbess and a local surgeon, every Saturday, singly examined women ‘in the home’ and if any of them had contracted any illness by their whores, they were separated from the rest and not allowed to prostitute themselves for fear the youth who had to do with them should catch their distempers.” This account is an early record of screening applied in the cause of community medicine and antedates by at least five centuries, the use of medical examination for the apparently healthy to prevent the spread of disease.

The precise origin of PHE is difficult to trace but it has been suggested that the intellectual beginnings were due to a British physician, Horace Dobell, a renowned clinician, author and expert on tuberculosis and diseases of the chest. Dobell proposed the periodic health examination as a way to identify “these earliest invasive periods of defect in the physiological state and to adopt measures for their remedy”.

In the United States (US), the public health use of screening probably first became established in the mid 19th century when, in conjunction with quarantine regulations, it was applied to immigrants. Its value in checking the flow of epidemics was taken for granted and at no time was its effectiveness tested. At the beginning of the 20th century, medical examinations were recognised as being of use to insurance companies for the purpose of rejecting or loading the policies of poor risk clients. The terms PHE and screening have similar meanings and are used interchangeably.

PHE in the US

In the US, the first universal PHE proposals began to appear at the turn of the 20th century. Possibly the first to suggest this was George Gould, a national figure in the medical community. In 1915, the National Tuberculosis Association designated a week for general physical examinations. This helped popularise these examinations as a tool for the early diagnosis of disease in general, and tuberculosis in particular.

Organised medicine also played a major role in the development of the PHE. In 1922 the American Medical Association (AMA) officially endorsed PHE and began a campaign to spread its practice in 1923. A manual for physicians was published. George Rosen argued that organised medicine saw the examination as serving the instrumental purpose of enhancing the position of the practitioner in the community, particularly in the wake of its opposition to compulsory health insurance. However, it has to be noted, that in the early 20th century, there was a great deal of neglect of periodic health examinations and apathy from the public.

The benefits of screening were first demonstrated by the use of mass miniature radiography (MMR) for the identification of individuals with tuberculosis. The use of MMR became common in many countries with the introduction of effective treatment for tuberculosis after 1946. With the reduction in the burden of tuberculosis, the application of screening for other chronic conditions began to be considered. This was particularly marked in the US, where a law on the control of chronic disease and the availability of screening was passed in the late 1950s. A Commission on Chronic Illness was founded in 1957, and a major review published in the Journal of Chronic Disease.

There was also the development of what became known as “multi-phasic screening”; the performance of multiple tests aimed at detecting unrecognised disease or defects. The objective was that screening should involve physicians only minimally, and be done by using technology that could be applied economically and efficiently. First conducted in
California and Massachusetts, this spread throughout the US. Prepaid group practice health care was a final important influence on the periodic health examination throughout the mid 20th century.

The most influential experience was that of the Kaiser Permanente Health Plan in the San Francisco area, due to the financial incentive structure of prepayment. An objective was the satisfaction of overwhelming patient demand for health check-ups. This led Kaiser to seek ways to maximise the efficiency of the examinations. An automated multi-phasic screening procedure that incorporated computerised test equipment and data analysis was developed. Opinion polls indicated a growing popular belief in the value of these examinations. However, it began to be accepted that the effectiveness should be examined.

**PHE in the UK**

The situation in the UK was somewhat different. In 1926, an experiment was established in London at the Peckham Pioneer Health Centre. This provided not only general practice and medical care, but also family planning advice and routine family health screening. This screening consisted both of formal medical examinations with laboratory tests and informal surveillance. While the purpose of this screening was not simply to act as a disease sieve, it was a natural, though incidental, part of the general aim of promoting health and attending the health centre. Nothing short of a periodic health overhaul on a national scale, the health centre thought could lead to the rational application of medical science and the elimination of sickness! When the National Health Service was established in 1948, the lessons learned from the Peckham experiment were ignored. Unlike the US there was no comparable demand for PHE.

**PHE in other countries and WHO guidance**

Experiments and practice in PHE were also undertaken in Australia, Canada, France, Germany, Ireland, Israel, Italy, Japan, New Zealand, Norway, Scotland and South Africa. Despite the enthusiasm, or perhaps because of it, in the late 1960s, the value of screening (periodic health examinations) began to be tested and examined more critically. WHO commissioned a comprehensive review of screening worldwide which enunciated ten common sense criteria to be applied before consideration of screening for a particular disease. These were:

1. The condition sought should be an important health problem.
2. There should be an accepted treatment for patients with recognised disease.
3. Facilities for diagnosis and treatment should be available.
4. There should be a recognisable latent and early symptomatic stage.
5. There should be a suitable test or examination.
6. The test should be acceptable to the population.
7. The natural history of the condition including development from latent to declared disease, should be adequately understood.
8. There should be an agreed policy on whom to treat as patients.
9. The cost of case-finding (including diagnosis and treatment of patients diagnosed) should be economically balanced in relation to possible expenditure on medical care as a whole.
10. Case-finding should be a continuing process and not a 'once and for all' project.

**Definitions and objectives**

In 1968, the Nuffield Provincial Hospitals Trust produced a book which dealt with the current state of the art of screening for a number of conditions. The general tone of this book was less enthusiastic than the literature of the preceding century. A series of articles in the *Lancet* was also more critical of screening and it emerged that most of the good evidence in favour of screening centred around pregnancy and early childhood. Following this series of articles, letters to the *Lancet* revealed elements both of confusion and dissent in the debate on whether public health services should increase their involvement with screening. Sackett and Holland, in characterising the opposing teams as ‘snails’ and ‘evangelists’, produced a cogent explanation of the main elements in the debate. They observed that the key to much of the argument lay in confusion over the use of terms, particularly what was meant by screening or PHE. There is still no universally accepted definition, and most doctors simply look upon screening (or PHE) as the use of any sort of test to identify possible disease.

Sackett discussed motives for screening, suggesting that there were four reasons. Firstly, to influence the gamble of life insurance; secondly, to protect people other than the patient, as in industrial and public health screening; thirdly, to obtain clinical baselines and fourthly, to do the patient some good, so called, prescriptive screening. Obviously there are also other motives, such as financial reward and biological research, satisfying public and medical demand and gaining information for administrative purposes. Many times these are all combined in one.

**Effectiveness of screening and PHE**

One of the earliest attempts to demonstrate the effects of screening originated with the use of screening by insurers. Knight produced evidence that such screening might be effective in saving lives. He reported that over a five year period, only 217 deaths occurred, where 303 would have been expected in an uninsured population of 6,000. A similar study on a population of 20,648 men having employer sponsored PHEs reported a favourable ratio of actual to expected deaths, but was more cautious in interpretation and recommended that prospective studies should examine the question.

There are at least three ways in which false conclusions can be reached in this field. Firstly, regression towards the mean, a natural tendency for high or low readings on one occasion subsequently to be nearer their mean level, may easily, and erroneously, be interpreted as clinical improvement in longitudinal screening follow-up studies. Secondly, the increased survival, of say, a cancer patient after detection by screening might be interpreted as a benefit of early diagnosis but in fact, be a reflection of the so-called ‘lead’ time, that is that the total duration of the disease process has itself remained unchanged, but since the diagnosis was made earlier, the patient appears to survive longer. Finally, there is the danger that the net of any screening process will tend to select out the more chronic and least severe diseases which by definition, will have a more favourable, clinical cause - again, tending to suggest that screening was beneficial.

Rodney Beard in 1959, an influential voice in the US, suggested that periodic medical examinations should be examined more critically. He pointed out that there were two factors in PHE: disease detection and health counselling. The randomised con-
trolled trial undertaken at the Kaiser Permanente was one of the first to examine the problems of periodic medical examinations. About 10,000 people were chosen at random from the 46,000 who were Plan members in 1964, and they were then randomly divided into equal control and treatment groups. The first seven years of the study were evaluated and comparisons made. Overall death rates were not significantly improved in the screening group. However, two specific causes of mortality, in particular age groups, did appear to be significantly improved in the screening group. But as sixty significance tests were performed on these mortality data, this is approximately the same outcome that one would expect purely by chance. The authors failed to make this point clearly in the discussion of their results. Morbidity measures, such as physician consultation rates and hospital admission rates did not significantly differ between the groups. There was little change in these findings at sixteen years.

A further controlled study was undertaken by Olsen, Kane and Procter, on smaller numbers and over only three years. This failed to show any measurable morbidity benefit in favour of screening.

A well designed controlled trial of screening was carried out in Malmö, Sweden, in 1970. Men born in 1914 and residing in the town were randomly allocated to screening and control groups. Following screening, intervention concentrated upon treating blood pressure greater than 165/110 and smoking. After four years the death rate in the two groups was not statistically significantly different, however, there was a significant shift in the causes of death. Twice as many men died of cardiovascular disease in the control group. This was offset by nearly twice as many deaths from cancer and other causes in the screening group. As the authors did not provide any evidence that they had lowered blood pressure and smoking levels in their screening population over the four year period, it might well be that the observed difference in their screening population over the four year period was due to random fluctuation rather than the screening programme.

A further randomised controlled trial was undertaken in England by the Department of Clinical Epidemiology and Social Medicine at St. Thomas’ Hospital and group general practices in St Paul’s Cray in Kent. Within two large group practices, all persons aged 40-64 in 1967 were identified and then randomly allocated, by family and within a general practitioner list, into two equal groups, designated ‘control’ and ‘screening’. The screening group, numbering 3,297, was invited by personal letter from their general practitioner to be screened.

The overall response rate of those attending for screening was 73.4%. Subsequent to screening, all information was passed to the general practitioner who then did a physical examination on each subject and decided on further tests, diagnosis and treatment. The same group of patients were invited to re-attend the screening clinic in 1969, the response this time was somewhat lower being 65.6%. Both the control and screening groups were examined after seven years and their levels of function were assessed. Overall, the mortality in the screening and control population was not significantly different. No significant difference appeared between the study and control groups for any of the various causes of death. There were no significant differences in certified sickness absence, use of home help or hospitalisation between the control and screening populations. There appeared to be a higher overall consultation rate observed in the screening population compared to the control, but that could perhaps be because of the need to investigate the findings at the initial screen. The economics of screening showed that, if introduced in the total population, it could increase costs to the National Health Service by about 10%, largely because of the examinations which had to be done following screening. Thus, the value of periodic health examinations (or screening), in this population in the UK was doubtful.

In Japan, comprehensive periodic health examinations have been undertaken for many years. There is little evidence of any, other than belief, that it is of any value. Analysis of the procedure showed that it increased health care utilisation and costs. The only possible benefit was that there was an increase in health care utilisation.

A more recent systematic review of the value of periodic health evaluation in 2007 was unable to show any improvement in outcomes in any of the studies examined. The authors concluded that PHE improves delivery of some recommended preventive services and may lessen patient worry, although additional research is needed to clarify the long term benefits, harms and costs of receiving the PHE. This was not a very comprehensive review as two controlled trials were omitted and the assessment of the methods and results of others, was not very critical (for example, 18, 22).

Changes in attitude in recent years

Opinion has changed in the United Kingdom over the past twenty years. Whereas in the 1960s, PHE’s were promoted by the medical profession and disregarded by the population, now the medical profession considers them to be of dubious value and view them with scepticism.

A recent article in the consumer magazine Which? cautions the population on the dubious value and view them with scepticism.

The aims of health policy – improvements in health – recent development

Health measures are intended to improve health. The aim of screening and PHE is intended to identify symptoms and signs in individuals at a stage when the condition is treatable and reversible. An alternative for health improvement and reduction in the burden of common diseases which have their roots in lifestyle, social factors and the environment depends upon a population-based strategy of prevention rather than an approach to identifying and treating ‘high-risk’ individuals. These alternatives have been analysed by Rose who concludes that a population-based approach is better and more cost-effective for common diseases. For relatively
uncommon conditions screening may be better if an effective form of treatment is available. There have been many publications reviewing individual screening tests (for example Holland and Stewart). Some health administrations have established national bodies responsible for reviewing individual tests and procedures (for example, UK National Screening Committee, US Preventative Services Task Force).

In spite of all these reservations identified by research, the UK government is planning to implement ‘Health Checks’ in England for all individuals aged forty and over. These proposals have not been included in the procedures approved by the UK National Screening Committee since they do not meet its stringent criteria. The term ‘health check’ is also a misnomer, what is intended is a structured risk assessment in general practice for coronary heart disease, cerebrovascular disease and type II diabetes. Economic assessment has been done on the basis of a number of complex models. The modelling shows that the impact of the programme would be significantly beneficial. The total cost per ‘Quality Adjusted Life Year’ is estimated at about £3,500, which is considered to be a very good use of NHS funds by the government.

Periodic health checks and screening are now very popular procedures. A priori they should be effective in reducing the burden of disease and improving well-being. Unfortunately the reality seems to be less attractive. Most of the proponents neglect the unfortunate side-effects of increasing anxiety, overtreatment and overdiagnosis which have been well-documented. Some of the randomised controlled trials which assessed outcomes such as mortality, morbidity and disability have not confirmed the hopes of the well-intentioned proponents, in spite of changes in some of the risk behaviours and increase of health service utilisation, particularly in countries with a competent primary care system. It is unfortunate that health authorities, worldwide, have neglected the need to conduct rigorous, long-term, pragmatic controlled trials with appropriate outcome measures and rely on superficially attractive policies based on modelling and ‘expert opinions’ or process measures. This short-termism is an unfortunate manifestation of all current policies – whether it be economic or health. The divide between snails and evangelists persists. It is unfortunate that the lesson propounded by Rose is neglected in the UK. Cardiovascular disease (CVD) (and risk factors) is extremely common in the UK population and part of health policy is to reduce inequalities. Common diseases have their roots in life style, social factors and the environment. Successful improvements in health must be based on population strategies. It is far more effective to change the population mean levels of risk (such as smoking habits or cholesterol level) rather than to tackle individuals with high risk levels, as is envisaged in the concept of ‘vascular checks’. Population strategies are a far less attractive public relations option and are more difficult politically, but would have a more profound effect in improving the levels of health of the UK population.

In the most recent document for the prevention of CVD at population level issued by the National Institute for Health and Clinical Excellence (NICE) the draft guidance states “interventions focused on individuals have tended to dominate CVD prevention activities. However, the largest overall benefit could be achieved by making changes (albeit small ones) within the population as a whole. As indicated by the Rose hypothesis, a small reduction in risk among a large number of people may prevent many more cases, rather than treating a small number at higher risk. A whole-population approach explicitly focuses on changing everyone’s exposure to risk. This may be best achieved through ‘upstream’ interventions: fiscal measures (including taxation), national or regional policy and legislation (including, for example, legislation on smokefree public places or the way food is produced).

Social and economic action can also result in a change in CVD risk (in such cases, the health outcomes are side effects – albeit desirable). Voluntary action may be effective. Sometimes, however, it may need to be supported by mandatory measures, for instance, when the pace of change is insufficient. Data from ‘natural experiments’ in a whole population (where there were no randomised controlled trials to assess the results) provide compelling evidence. One example is the reduction in the consumption of animal fats in Eastern Europe, following the break-up of the Soviet Union (29). Another example is the introduction of legislation in Mauritius to make it mandatory to use polynsaturated oils as a substitute for highly saturated cooking oils. In such cases, there has been a remarkably rapid reduction in CVD mortality among the populations. Conversely, rapid rises in CVD mortality have been seen in China and elsewhere, principally due to the adoption of a Western diet rich in saturated fats.

Interventions which rely on people deciding to change their behaviour are likely to vary in effectiveness. For example, people who are disadvantaged might find it more difficult to change than affluent people. As a result, some interventions that focus on changing behaviour may inadvertently increase health inequalities. To overcome this, the recommendations do not, in the main, rely on individual choice but, rather, aim to make the healthy choice the easy choice. Hence, the emphasis is on changing policies, systems, regulations and other similar ‘upstream’ factors. This approach is likely to reduce, rather than increase, health inequalities and is congruent with NICE’s guidance on behaviour change.

The use of preventive health examinations or multi-phasic screening may be justified for populations that have few medical facilities, or organised health systems, to identify those individuals who require care or treatment. In most developed countries individuals have easy access to health facilities so that benefits to improve the population’s health through these measures are difficult to identify. The occasional individual may benefit – but at the cost of harm to others. It is for this reason that WHO, US, Canada, the UK and other health authorities have introduced clear principles by which screening measures should be assessed before introduction to medical practice. PHE is an attractive commercial undertaking leading to the consumption of drugs and the use of diagnostic equipment. The stake that industry and advertising have is exemplified by the unscrupulous behaviour towards those who are more questioning. It is also, superficially, very attractive to the layman, after all our cars have to be tested and examined regularly to ensure that the engine and brakes work. Thus health policy and politicians react favourably to such populist procedures – but completely neglects that all such examinations are fallible, leading to unnecessary further tests, the induction of anxiety and to being ‘labelled’ ill.

REFERENCES


---

Nordic Health Care Systems
Recent reforms and current policy challenges

Jon Magnussen, Karsten Vrangbaek and Richard B. Saltman

Published by Open University Press

339 pages

The Nordic model of health care systems is assumed to contain consistent features across all five Nordic countries: tax-based funding, publicly owned and operated hospitals, universal access based on residency and comprehensive coverage. The reality is considerably more complex, with great variation at the structural level in the way that institutions are designed and at the policy level in the way strategies are conceived and implemented.

This book examines recent patterns of health reform in Nordic health care systems, including the balance between stability and change in how these systems have developed. Detailed comparisons are undertaken along a variety of policy-driven parameters.

Freely available at: www.euro.who.int/observatory/Studies/20091021_2