

## TESTING A TEST

The evaluation of a treatment is often difficult, but at least there is a clear outcome - either the patient gets better or does not.

When assessing a new diagnostic test, however, the situation may not be so simple. The successful "outcome" for a test may be either a positive or a negative result, depending upon whether a doctor wishes to confirm or exclude a diagnosis. Another possible way of assessing outcome is by whether the test affects the management of the patient: some test results, both positive and negative, will lead to a change in management, whereas in other circumstances a positive or negative result will lead the doctor to continue with the same management plan.

## Diagnoses as drivers

The development of a new treatment is often seen as a very important driver of health service direction and costs. The development of new diagnostic tests is equally important. When a new diagnostic test is developed, for example the ability to identify an enzyme deficiency, then a new group of people will be identified as having a positive test result and this will initiate further diagnostic effort or treatment, even though there is not always good evidence that the treatment will be beneficial.

The introduction of commonly available tests to measure cholesterol has led to treatment costs of about £1.5 million per million population per year, even though our national policy is still not to screen the whole population for raised levels of cholesterol.

## Sensitivity and specificity

Diagnostic tests can be classified as having either positive or negative results. Even those tests that have a continuous distribution can be defined as having an arbitrary cut-off point distinguishing positive from negative. This is done, for example, in Down's syndrome screening where a single cut-off point is taken as the level at which further investigation is justified.

People who have tests done either have or do not have an underlying disease or condition, and the relationship between these debatables can be most elegantly expressed by a two-by-two box.

	TEST	
	Positive	Negative
Disease present	✓	X
Disease absent	X	✓

Using data from real clinical studies in this simple two-by-two box, it is possible to assess the sensitivity and the specificity of the test.

The sensitivity of a test is the proportion of people with the disease who have a positive test result. The higher the sensitivity, the greater the detection rate and the lower the false negative rate.

The specificity of the test is the proportion of people without the disease who have a negative test. The higher the specificity, the lower will be the false positive rate and the lower the proportion of people who have the disease who will be unnecessarily worried or exposed to unnecessary treatment.

## Predictive value

The positive predictive value of a test is the probability of a patient with a positive test actually having a disease. The negative predictive value is the probability of a patient with a negative test not having the disease.

While the sensitivity and specificity of a test are constant within the populations under test - and generally wherever the test is performed - the predictive value of a test result depends not only on the sensitivity of the test but also on the prevalence of the condition within the population being tested.

For instance, if a test has a sensitivity of 96%, then the positive predictive value varies with the prevalence as follows:-

Prevalence (%)	PPV (%)
1	19
2	33
5	56
10	73
20	86

The significance of a positive test may be very different in hospital than in general practice because the prevalence of disease is higher in the referred population. Hospital doctors frequently criticise general practitioners for "missing a diagnosis" because the significance of a positive test result is, for most tests, statistically different in hospital practice than in general practice.

## The clinician's perspective

Clinicians judge tests primarily with respect to sensitivity and specificity, often underestimating the importance of

predictive value when transferring the test from secondary to primary care.

However, the clinician does not necessarily seek a test with a high sensitivity and a high specificity.

- When it is very important not to miss a diagnosis, a test with high sensitivity is chosen.
- When it is very important not to create false positives, because of the serious consequences of a positive diagnosis, the clinician will give greater emphasis to assessing the test with respect to its specificity and will seek a test with higher specificity, even though there may have been a trade-off in the test with lower sensitivity than the alternative.

## The purchaser's perspective

---

The purchaser's perspective obviously has to take into account the clinician's perspective, but the purchaser is likely to have different criteria for assessing a new test. These will include:-

- What is the predictive value of the test in secondary care?
- What is the predictive value of the test in primary care?
- What impact will the test have on the quality of care offered? Can it, for example, replace a more painful or distressing test?
- What are the adverse effects of the test?
- What is the cost per test?
- What is the cost per case diagnosed, including the cost of investigating or treating false positives?
- What would be the impact of the test on the total cost of care?

## Commissioning tests

---

Although district health authorities are having a growing influence on treatment and care services through their power to purchase and specify what services their population needs, they have relatively little influence over diagnostic services, even though those services have a major impact on their population and health care costs. GP fundholders have an influence on both imaging and laboratory services but the commissioning of diagnostic services is principally carried out by clinicians, a fact recognised in many Trusts by the move to set up an internal market with the budget for imaging and laboratory services being allocated to the clinicians who use those services. The choice of diagnostic services is therefore made by clinicians with the costs of diagnostic services being incorporated into the price of the treatment services which health authorities commission.

The approval of business and strategic plans of Trusts will offer the Regional Office an opportunity to identify and challenge a Trust business plan to make some major investment, for instance to buy a new CT scanner. However, as in so many areas of clinical practice, transformation comes

about not by a small number of major investments but by a large number of small developments, and the development of new patient testing make this aspect of technology-creep one that will be even more difficult to control in future.

## Evidence-based diagnosis

---

New diagnostic tests can be subject to rigorous evaluation and steps can be taken to appraise the evidence on which the decision to introduce a new diagnostic test is taken, but this approach is used all too rarely.

Diagnosis lies at the heart of clinical practice and developments in diagnostic services drive changes in health care. The direction and intensity of this drive is presently determined by those who invent, make or market diagnostic equipment, and by the clinicians who use new diagnostic opportunities. If best value for money is to be achieved, those who pay for health services must examine the evidence of every new diagnostic test much more rigorously.

---

---

## STROKE REHABILITATION

Stroke represents a condition which is responsible for the consumption of a large proportion of the NHS budget; about 4% is spent each year on cerebrovascular disease, and the biggest part of that is on the aftermath of stroke.

### The size of the problem

---

A district with a population of 250,000 will have 500 first-ever and 100 recurrent stroke cases each year (in total about 0.25% of the population). Most strokes occur in over-65s, and about 20% of stroke victims die in the first four weeks, with a further 10% dying within a year. Stroke deaths account for 12% of deaths from all causes in England and Wales.

The typical district of 250,000 is likely to devote about 30 beds to the care of stroke victims who have recently had a stroke, (about 12% of beds on general medical wards).

The typical district will have around 1500 stroke survivors living in the community, of which half (750, 0.3% of the population) will have a significant level of disability. Half of stroke survivors have some significant disability.

### Effective rehabilitation?

---

One important question is whether rehabilitation after stroke is effective. This has been well reviewed in an *Effectiveness Bulletin* from the University of York in March 1992. This bulletin examined 17 major studies; important lessons could be drawn, despite their comment that there were few well designed studies that assess the effectiveness of rehabilitation after stroke.

## Organisation saves lives

---

Organised stroke care can, however, save lives. Langhorne and colleagues published a statistical overview of randomised controlled studies reported between 1962 and 1993 in which the management of stroke patients in a specialist unit was compared with that in general wards. The overview covered 1568 stroke patients, with 766 in special units and 820 in general wards.

The conclusions were that patients treated in stroke units showed a reduction in mortality of 28% after four months, and 21% over one year.

Death is an objective outcome: morbidity is less so. A detailed analysis of functional outcomes was hampered by lack of consistent recording. However, nine of the ten studies reported functional gains in stroke units, and one reported very similar outcomes.

## Northampton's GRiP

---

In Northamptonshire, as part of the GRiP initiative, they have a stroke project whose basic aim is to answer the question "can we put evidence derived from research into clinical practice". While research evidence often provides provision of the best possible care, this has to be viewed in relation to the resources available to produce a pragmatic solution which is acceptable and achievable at the local level.

## Specialty Liaison Groups (SLG)

---

Services for stroke survivors are provided by many professionals and agencies, whose behaviour may need to be changed. To have an impact on behaviour change, the 'target audience' needs to be receptive, and the message needs to be timely and credible.

Northamptonshire is developing a particular interest and expertise in the use of small groups of peers spanning primary and secondary care. The SLG system, where hospital consultants talk directly to GPs and other professionals, with the involvement of DHA purchasers and public health, can act as a lever for change of professional practice when driven by research evidence.

Clinical guidelines (including a referral protocol) are being designed to improve stroke care in hospitals and in the community, by recommending good clinical practice based on research evidence, and with an emphasis on the primary/secondary care interface, and integrating rehabilitation services.

## Credibility

---

Unless credibility has been established in the minds of clinicians, then the prospect of useful change is poor.

Northamptonshire tackled this issue by commissioning its own detailed review of stroke rehabilitation, clearly presented.

Northamptonshire held its own local Stroke Conference, with local and guest speakers, and with attendance across professional groups.

Northamptonshire decided that the project would have few administrative meetings, but with clear agendas and more action between meetings.

## Audit

---

The main levels for change are the quality of information between clinicians in provider organisations and purchaser HAs, and the clinical audit.

Retrospective audits formed the prior baseline to the project, with ongoing audit to monitor effectiveness of change on an ongoing basis.

## Clinical guidelines and contracts

---

The steering group is examining brief checklists or 'pathways', backed up with short reference documents summarising diagnostic criteria, treatment options, indications for referral and discharge, arrangements for follow up, information for patients and points for audit. Evaluation forms will also be designed with each guideline.

These guidelines are being developed in a process that entails participation by key groups in a full consultation process, maintaining credibility and 'ownership' of the process.

Guidelines will, of course, be open to update when new evidence is available, or based on the experience of operating the system. Guidelines will form a basis for future contracts.

## Effectiveness in action

---

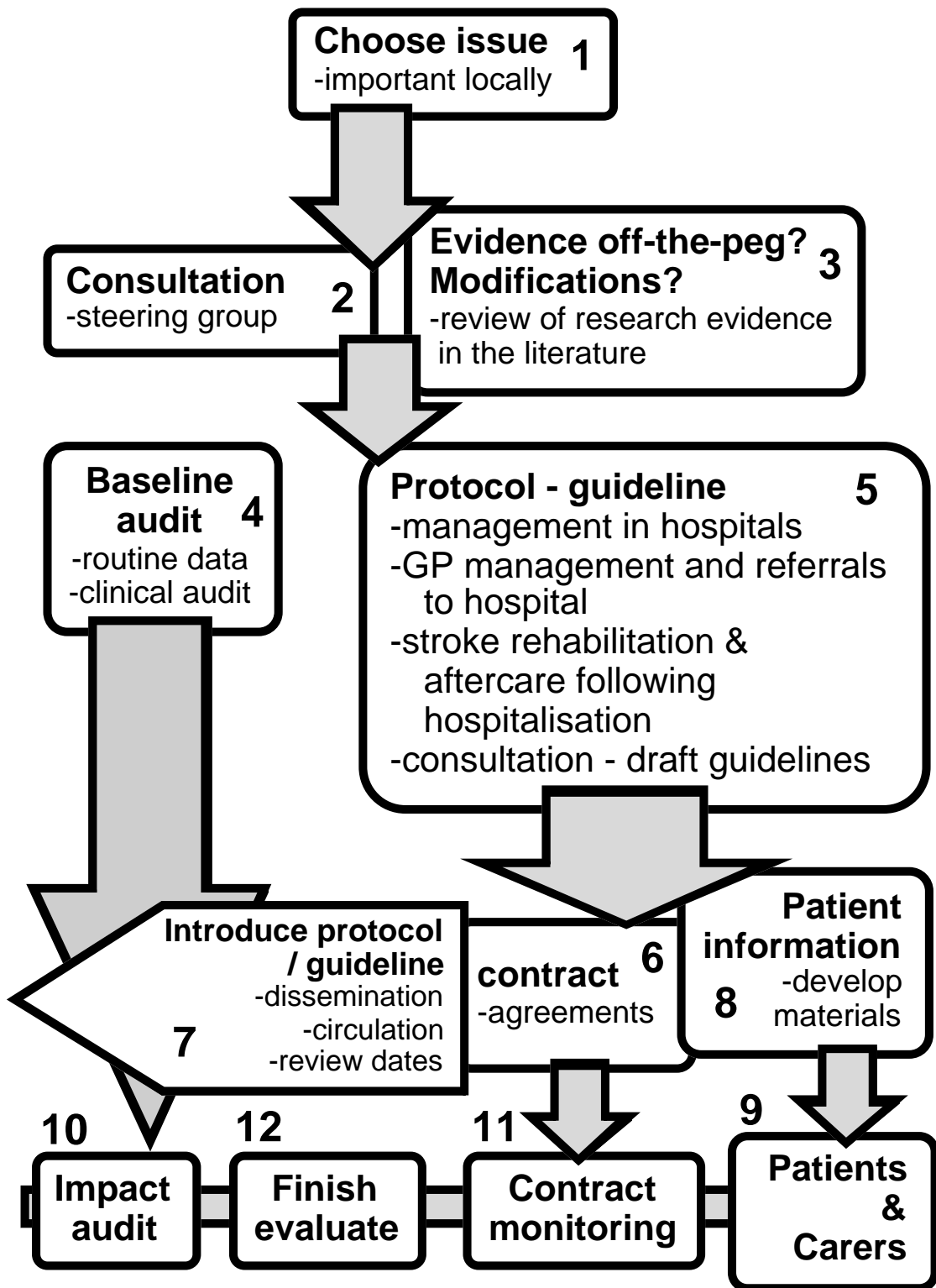
The Northamptonshire GRiP project on stroke rehabilitation is drawing on an extraordinary breadth of experience. Public health, consultants, GPs, therapy services, social services, community nursing, purchasers and finance are all represented on the steering group.

The project is based on effectiveness, and its aim is to get that evidence of effectiveness into action.

## References:

Effectiveness Bulletin on Stroke Rehabilitation,  
University of York, March 1992.  
Langhorne et al, Lancet 1993 342: 395-8.

# Northamptonshire GRiP: Stroke Care



# DERMATOMYOSITIS

## Background

Dermatomyositis occurs infrequently, with an incidence probably less than 1 per 300,000 per year. It is characterised by an intramuscular microangiopathy mediated by a complement membrane attack complex. There is a loss of muscle capillaries, muscle ischaemia, muscle-fibre necrosis and perifascicular atrophy.

There is an unmistakable rash with purplish appearance on the eyelids, cheeks and light exposed areas. A rash is often seen on the extensor surfaces of knees, elbows and knuckles, and this becomes scaly.

Treatment is usually with steroids, but cytotoxic drugs and immunosuppressants are also added. Many patients do not respond to this treatment and remain physically disabled.

## RCT shows startling results

A very well conducted double-blind placebo-controlled randomised control trial of high dose immune globulin has shown that it is an effective treatment. The study from the NIH in Bethesda looked at fifteen patients with therapy resistant dermatomyositis. Patients were block randomised to immune globulin infusion or placebo by disease severity, and they continued on their normal therapy during the study. Immune globulin was infused at 2g/kg body weight, with one infusion each month.

Treatment was for three months, followed by cross-over after a one month washout. Patients were followed for three months after the end of the study before codes were broken.

Response was assessed by neuromuscular symptom scale, the ADL scale, a scale to measure muscle strength and repeated muscle biopsies.

The results were excellent. Of 12 patients treated with immune globulin, 9 had major improvements, 2 had mild improvement and only one had no change. Placebo had no such effect. No unwanted effects were reported with immune globulin infusion.

Muscle biopsies after immune globulin treatment showed an increase in the diameter of muscle fibres, and regenerating muscle fibres became sparse. There was an increase in the number of capillaries and decrease in their diameter, and the ratio of muscle fibres to capillaries returned almost to normal.

## Patients back to normal

Perhaps the simplest way of conveying a complex set of objective scale results is to quote from the paper: speaking of five patients with severe muscle weakness the authors said that "these patients had not felt so strong since the onset

of their disease ....Those using wheelchairs were able to get up, run, climb stairs and behave normally."

After the study patients actively sought to obtain continued immune globulin treatment. Six managed to obtain it, and require an infusion approximately every six weeks.

High dose infused immune globulin was effective in treating therapy-resistant dermatomyositis. Patients who were previously disabled returned to normal life.

## Cost Implications

Treatment for an individual patient at 2g immune globulin every six weeks would be expensive - approaching £20,000 per year. There are very few patients, however. That can, however, be balanced by reduction in other drugs, some of which are expensive. In addition, the very significant reduction in disability in these patients, enabling many to lead normal lives, would have concomitant economic benefits.

## Reference:

Dalakas et al. New England Journal of Medicine 1993 329: 1993-2000.

## Questions to be Answered

- Q: What need is met by this treatment?  
A: Disabled patients with dermatomyositis can be returned to normal life.
- Q: What happens at present?  
A: Patients are treated with steroids, immunosuppressant and cytotoxic drugs. Many do not respond.
- Q: How does the treatment improve effectiveness?  
A: Patients resistant to standard therapy can achieve normal life. Intravenous immune globulin therapy can be given at home.
- Q: What is the capital cost?  
A: None.
- Q: What is the treatment cost per case?  
A: About £20,000 per year.
- Q: Will this increase the total cost of care?  
A: Not known. There is a balance between high cost of immune globulin treatment, to be offset against reduced cytotoxic and immunosuppressant drugs, as well as reduced care costs and the economic benefits of returning a disabled patient to health.

## Advice to purchasers

- 1 Will increase quality and effectiveness.
- 2 Will probably increase total costs.
- 3 Should consider inclusion in specification for intravenous immune globulins for as part of longer term development of clinical immunology services.

# OSTEOPOROSIS

The first Effectiveness Bulletin from the University of York concerned screening for osteoporosis to prevent fractures. It examined the question whether population based screening programmes should be established to prevent fractures in elderly women. It assumed a model of identifying a high-risk group by screening the whole population at the time of the menopause.

## The size of the problem

---

Fractures in post-menopausal women are an important cause of morbidity, mortality and cost.

- During any five-year period, 10% of a population of women aged 70 years and above will suffer a hip fracture. 10-20% of these will die as a result.
- The average length of hospital stay is 23 days.
- Patients with hip fracture account for over 20% of all orthopaedic beds.
- The average cost of hospital stay is £2500.
- Only one third of survivors are fully mobile after six months.

The incidence of hip fractures in women is age-related. From an incidence of about 0.5% in the range 70-74 years, it rises to nearly 1.5% a decade later and is very high in women over 85 years.

## Is there an effective treatment?

---

There are no simple treatments for established osteoporosis. Treatment is therefore aimed at the time of rapid bone loss during the menopause. HRT based on oestrogen, alone or in combination with progesterone, has been shown to retard, stop or even reverse bone loss after the menopause.

HRT is recommended for a maximum of 10 years. That leaves a treatment gap of about 15 years between stopping HRT and the age at which fractures become common. HRT has been shown to reduce fracture incidence by about 50% in relatively young women, and that probably overestimates the protective effects of HRT in elderly women.

It is not clear how long the HRT effect persists. There is evidence that the treatment effect diminishes after treatment stops, and within a few years the protective effect may have worn off.

Estimating the effectiveness of HRT on reducing fractures in elderly women is impossible.

## Can high-risk patients be identified?

---

The standard test is for bone density. The sensitivity and specificity of bone density measurements in identifying those women who will go on to have fractures later in life is not established.

If the 20% of women with lowest bone density measurements are taken as the high risk group, then only 28% of those would have gone on to suffer fractures later in life in the absence of therapy. Women with bone densities above this cut-off will suffer 63% of all fractures.

It is possible that biochemical measures to identify rapid bone losers would be more effective, and some new assays are becoming available, but have yet to be fully evaluated.

## Will women come for screening?

---

It is expected that even with a lot of effort only 70% of women will take up screening opportunities.

## Will women accept long-term HRT?

---

Long-term compliance with HRT is as low as 30%.

## What is the overall impact?

---

It is likely that a screening programme using bone density measurement and long-term HRT in the high risk group will prevent fewer than 3% of fractures in elderly women.

## Implications for Health Authorities & GPFHs

---

This is a detailed and well reasoned review, as well as being a good read. It has assembled a solid body of evidence, and is a paradigm for anyone considering a screening programme.

Effective Health care Bulletins are available from School of Public Health, University of York, 32 Hyde Terrace, Leeds LS2 9LN. Price £3 or £25 for a series of nine bulletins.

---

---

## NEAR-PATIENT TESTING

The Health Services Research Unit at the University of Warwick has produced a report on near-patient testing (NPT) in general practice. The objectives were:-

- To investigate the ways in which GPs make use of the tests marketed for NPT when they are made available in the surgery.
- To investigate the impact on patient care of two selected tests.
- To investigate the economic aspects of NPT in primary care.

The research involved a four month baseline which involved observation of current laboratory use for all tests. This was followed by a 12 month period during which practices were given access to bacteriological and biochemical NPT for six months each, when both laboratory and NPT use were

monitored. Finally practices retained the NPT equipment for nine months, during which time a detailed study was made of the impact on patient care of one biochemistry and one bacteriology test.

The NPT tests examined were:-

## **Bacteriology**

MSU for UTI.  
Chlamydia

## **Biochemistry**

Cholesterol  
Haemoglobin  
Gamma GT  
Sodium & Potassium

The biochemistry tests involved the manipulation of samples and instruments to obtain the results, as did some of the bacteriology tests.

A total of 25,300 tests was analysed, with test costs estimated in practice and laboratory settings. Patient questionnaires were also used to measure extra associated costs.

## **Attitudes**

Most GPs were very keen about the potential of NPT at the start of the study. At the end of the study, this positive attitude had largely evaporated. Concern about time, both their own and practice nurses was a major factor. Time involved in maintenance, quality assessment and time needed to perform the test were the major negative factors.

Practice nurses were more cautious initially, with a significant minority concerned about the time and work involved running the equipment. At the end of the study, one test (MSU for UTI) received widespread enthusiastic responses from practice nurses. Time pressure was felt as a major negative factor on performing NPTs by over half the nurses.

## **Quality assessment**

Considerable quality assessment was involved in the study, including internal quality control, parallel testing by local laboratories and external QA in local and national schemes. Some of the equipment needed checks carried out each day, and there was no doubt that this all contributed significantly to the negative impact of NPT.

Performance of the NPT testing in QA schemes was generally acceptable to good.

However, during the study only one case (out of 25,300) came to light where a possible erroneous result may have led to incorrect diagnosis or treatment.

## **Uptake of NPT**

The practices observed averaged only 9 NPT tests per week.

There was much greater uptake where there was enthusiasm from a practice nurse, who was in control of equipment and testing.

Some tests increased markedly, especially cholesterol testing, while there was little impact on testing rates for other analytes.

## **Cost implications**

Because the average uptake of NPT was low, the cost per NPT test was high, and generally higher than laboratory costs. The cost per test could approach laboratory costs if there was a higher uptake.

Costs in this study included equipment costs, costs for the reagents or consumables required for the tests and controls, and patient costs. The major contribution was the consumable/disposable cost.

For cholesterol, the analysis showed that NPT could increase the average practice investigation costs by about £2,400, even with no increase in testing.

For MSU investigations, by contrast, the cost impact on the average practice would be negligible.

## **Patient management**

There was a positive impact only for NPT urine testing. For patients who received only an NPT test with no immediate laboratory follow up, there was a significantly higher level of antibiotic prescribing. There were also significantly fewer referrals and investigations of the urinary tract in NPT tested patients in the six month follow up period.

These positive benefits were seen only for NPT testing of MSU for UTI. There was no gain for patients for NPT cholesterol testing.

## **Cost effectiveness**

For MSU the observed reduction in referrals and investigations of the urinary tract for patients who first received a NPT test leads to sizeable potential annual savings of between £1,150 and £2,450.

Potential national costs for widespread national introduction of NPT for cholesterol and MSU for UTI were calculated. For cholesterol, a large health care cost increase of £16.5 - £24.5 million per year was estimated. For MSU for UTI, the savings predicted were £1.1 - £3.1 million.

## Survey conclusion

---

Widespread introduction of NPT into general practice is likely to have far reaching and complex effects. The paper from Warwick looks at these in detail, together with the policy implications that stem from the study. They particularly point out that careful attention should be given to which tests are required, for what purpose, and what specification, if NPT is going to be of future value in primary care.

## Bandolier's comments

---

One of the problems with the introduction of near patient testing is that manufacturers are making small versions of laboratory equipment for introduction into primary care. These units need much the same attention and sample handling as does true laboratory equipment, and the result is predictable - too much time demanded of busy practice staff, and with high unit test costs making NPT largely uneconomic.

For NPT to be successful, a different approach needs to be followed. Near patient tests that confer substantial benefit to GP and patient need to involve the minimum of sample handling, have a minimum of no reagent or disposable cost, and have results which are available virtually instantaneously.

Only when these criteria are met is it likely that the benefits so long trumpeted for NPT will be fulfilled.

## Details of useful publications:-

---

Near Patient Testing in General Practice costs £4 and is available from The Secretary, Health Services Research Unit, Warwick Business School, University of Warwick, Coventry CV4 7AL.

The Association of Clinical Biochemists also produces a helpful booklet entitled "Guidelines for Implementation of Near Patient Testing". It is available from ACB, Royal Society of Chemistry, Burlington House, Piccadilly, London W1V 0BN.

## PROSTATE CANCER: SCREENING, DIAGNOSIS, AND MANAGEMENT

Prostate cancer is a major public health issue in the United States and European countries. Recent correspondence in the Lancet last year turned up the heat in Britain as well. A number of studies are underway examining screening for prostate cancer.

The combination of an ageing population and greater awareness has led to increases in the number of these cancers being diagnosed. In the US in 1993 about 165,000 new cases of prostate cancer were diagnosed and 35,000 men died from the disease. There have been estimates that by the year 2000, the number of prostate cancer cases will increase by 90%, and the mortality by 37%. A 50-year old American man has a 40% chance of developing microscopic prostate cancer, a 10% chance of being diagnosed, and a 2-3% chance of dying from the disease.

This makes an up-to-date review quite important, and Marc Garnick provides that in *Annals of Internal Medicine*. The data source was a MEDLINE search of articles relating to the diagnosis, staging, screening, surgery, radiation therapy, medical management and research in prostate cancer. Results from long term studies (>5 years) are reported. A total of 139 references are included, up to about mid-1993.

This is an excellent review, and brings the reader up-to-date with developments. If it has a failing, it is perhaps that because of its breadth it lacks a little in depth.

## Other useful papers

---

Two other, slightly earlier reviews by the practising urologist and active researcher, Joseph Oertling of the Mayo Clinic, on the role of the effective tumour marker, prostate specific antigen (PSA), are useful in adding flesh to the bones. Much urological practice in prostate cancer treatment is now being driven by reference to tumour markers, and the Oertling papers cover the use of PSA in considerable detail, as well as reviewing in detail a significant number of the published accounts.

## References:

---

- Garnick. *Annals of Internal Medicine* 1993 118: 804-18.
- Oertling. *Current Opinion in Urology* 1992 2: 348-55.
- Oertling. *Journal of Urology* 1991 145: 907-23