failure: results from the SUPPORT project. Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments. *Circulation* 1998;98:648-55.
5 Shepardson LB, Youngner SJ, Speroff T, O'Brien RG, Smyth KA, Rosenthal GE, et al. Variation in the use of do-not-resuscitate orders in patients with stroke. *Arch Intern Med* 1997;157:1841-7.
6 Levin JR, Wenger NS, Ouslander JG, Zellman G, Schnelle JF, Buchanan JL, et al. Life-sustaining treatment decisions for nursing home residents: who discusses, who decides and what is decided? *J Am Geriatr Soc* 1999;47:82-7.
7 Sayers GM, Schofield I, Aziz M. An analysis of CPR decision-making by elderly patients. *J Med Ethics* 1997;23:207-12.
8 Shepardson LB, Youngner SJ, Speroff T, Rosenthal GE. Increased risk of death in patients with do-not-resuscitate orders. *Med Care* 1999;37:727-37.
9 Thompson BL, Lawson D, Croughan-Minihane M, Cooke M. Do patients' ethnic and social factors influence the use of do-not-resuscitate orders? *Ethnicity Dis* 1999;9:132-9.
10 Bowling, A. Ageism in cardiology. *BMJ* 1999;319:1353-5.
11 Royal College of Physicians. *Cardiological interventions in elderly patients. A report of a working group of the Royal College of Physicians.* London: RCP, 1991.
12 Ebrahim S. Demographic shift and medical training. *BMJ* 1999;319:1358-60.
13 Smith R. Medicine and the marginalised. *BMJ* 1999;319:1589-90.
14 Tonks A. Medicine must change to serve an ageing society. *BMJ* 1999;319:1450-1.

Valuing the effects of sildenafil in erectile dysfunction

Strong assumptions are required to generate a QALY value

Sildenafil is a true breakthrough drug in the sense that it provides a potential treatment for a condition for which there was no existing acceptable alternative. This complicates any attempt to describe the cost effectiveness of sildenafil in erectile dysfunction, such as that by Stolk et al in this week's *BMJ* (p 1165).[1] They compare sildenafil with papaverine-phentolamine injections, which they argue are rationed on "medical grounds" and will not achieve the population benefits that might be achieved by sildenafil. More controversially, they argue that "the incremental cost-effectiveness of sildenafil lies at the favourable end of the scale when compared with interventions in health care for other diseases."

The comparator therapy Stolk et al refer to, papaverine-phentolamine injections, seems not to have been rigorously evaluated and is not widely used. They used a cost utility approach in which a representative sample of the general population were asked to value the effects of treatment (for a condition that they did not have) to generate a cost per quality adjusted life year (QALY). Why might we question the validity of these findings?

Generating values for a treatment and comparing them with scores for other healthcare interventions requires a method for translating the clinical benefits attributable to treatments into a common metric, in this case the QALY. Stolk et al used a time trade off approach to transform benefits in quality of life to quality adjusted life years.[1] A population sample was asked to trade off the alternatives of being in a less desirable health state for a longer period, followed by death, versus being in a more desirable state for a shorter period followed by death.[2]

There are several well known assumptions, and many practical problems, associated with generating utility values.[3 4] Firstly, the QALY depends on an assumption that the trade off between different health states is known, rather than subject to uncertainty and measurement error. Secondly, a constant proportional trade off between risks is assumed—that is, we consider two years at a utility of 0.5 to be worth one year at a utility of 1 (perfect health). Thirdly, we assume that QALY valuations are independent of previous health states. Fourthly, many cost utility models are "black box" analyses where it is hard to disaggregate contributing components even when the methods are clearly written (as in Stolk et al's paper) so the validity of a model must be taken on trust to some extent. All these assumptions serve to question the validity of generating a single cost utility measure.

Papers p 1165

In evaluating the cost effectiveness of sildenafil, Stolk et al target the restoration of sexual function, and do not distinguish between the situation where there is one failed attempt at sexual intercourse in two attempts, or five failures in 10 attempts. In generating estimates of the utility of sildenafil Stolk et al did not take into account the experiences of men with erectile dysfunction in the trials, but based their estimates on a survey of the general population and thus on the imagination of their sample. The time trade off approach confounds time preferences with patient preferences, thus downgrading the importance of events that are in the distant future,[2] in this case death, which may make sildenafil appear a relatively valuable treatment when contrasted with a treatment for a condition where the threat to life is more immediate. The time trade off has also only moderate agreement with alternative methods for generating utility measures.[2 3]

The only convincing argument for conflating cost and utility information into a single summary measure (the QALY) is to compare treatments for a range of conditions. However, since there are so many good reasons to suppose that QALY estimates are derived using strong assumptions, are context specific, and are not comparable across different diseases, we may question whether Stolk et al's methods are the most appropriate.

Like many newly developed drugs, sildenafil is supported by a programme of randomised trials that provide good evidence of its clinical effectiveness. In the pivotal trials sildenafil was associated with a real improvement in sexual function.[5] A more robust cost effectiveness analysis might focus directly on the trial programme and provide estimates of the costs and effects attributable to sildenafil in the clinical outcomes measured—in other words, unpacking the black box and making explicit the costs and benefits of sildenafil. Estimates of usage and tolerability may be gleaned from the trial programme and open label extension studies. This approach will avoid the need for the

BMJ 2000;320:1156–7

strong assumptions required to fulfil the specification of the cost utility approach. Having established and described what the drug may achieve in use and at what cost, it is then a difficult political rather than technical decision whether it is made available.

In the United Kingdom uncertainty remains on whether the National Institute for Clinical Excellence will use QALY methods to redistribute resources for therapeutic interventions and diagnostic techniques.[6] The alternative is simply to assess each intervention on its merits and make recommendations on the basis of clinical effectiveness and cost considerations. When Professor Sir Michael Rawlins, chair of the National Institute for Clinical Excellence, commented that recommendations will be based on difficult judgments which have "no mathematical quantitative approach"[7] he appeared to be favouring the latter. This will more honestly reflect the evidence base, enable a broader public debate, and increase public accountability.

Nick Freemantle *reader in epidemiology and biostatistics*

Medicines Evaluation Group, Centre for Health Economics, University of York, York YO10 5DD
(meg@york.ac.uk)

NF has received funding for research from the Department of Health. The medicines evaluation group receives an unrestricted research grant from Pfizer Ltd, the manufacturers of sildenafil, for evaluating an intervention to promote improved care in diabetes.

1 Stolk EA, Busschbach JJV, Caffa M, Meuleman EJH, Rutten FFH. The cost effectiveness of sildenafil. *BMJ* 2000;320:1165-8.
2 Gold MR, Siegel JE, Russell LB, Weinstein MC. *Cost-effectiveness in health and medicine*. New York: Oxford University Press, 1996.
3 Dolan P, Gudex C, Kind P, Williams A. Valuing health states: a comparison of methods. *J Health Econ* 1996;15:109-31.
4 Carr-Hill RA. Assumptions of the QALY Procedure. *Soc Sci Med* 1989;29:469-77.
5 Goldstein I, Lue TF, Padma-Nathan H, Rosen RC, Steers WD, Wicker WA, et al. Oral sildenafil in the treatment of erectile dysfunction. *N Engl J Med* 1998;338:1397-404.
6 Freemantle N, Mason J. Not playing with a full DEC: why development and evaluation committee methods for the appraisal of new drugs may not be adequate. *BMJ* 1999;318:1480-2.
7 Yamey G. Chairman of NICE admits that its judgments are hard to defend. *BMJ* 1999;319:1222.

Multicentre research ethics committees: has the cure been worse than the disease?

No, but idiosyncracies and obstructions to good research must be removed

Papers pp 1179, 1182, 1183
Personal view p 1217

I first wrote about the byzantine labyrinth that surrounded obtaining ethics committees' approval for multicentre studies in England in 1995, as well as mentioning other unsatisfactory aspects of local research ethics committees.[1] At that time a working party of the chief medical officer suggested the establishment of multicentre research ethics committees on a regional basis to take care of multicentre studies. These were established in 1997.[2] So is it now simpler to obtain approval for multicentre studies? Are decisions reached more speedily? Are local research ethics committees restricting their comments on multicentre studies to local problems? Or has yet another layer of bureaucracy been added, making the process even more labyrinthine?

In the past two years frustrated research workers have regularly told me that the new system is a disaster. Early feedback suggested that local research ethics committees were finding their subordinate role difficult. These committees have always jealously guarded their independence**.** The early problems led to further guidance from the Department of Health and NHS Executive on the precise responsibility of local research ethics committees for these multicentre applications. Most importantly, the guidelines stressed the need for speed (a response within three weeks) and that objections should be based solely on local issues.

Has this worked? Two papers in this week's *BMJ* are highly revealing (pp 1179, 1182).[3 4] Both look at the fate of a multicentre study submitted to and approved by the appropriate multicentre research ethics committee. The study of Tully et al is the larger, involving 125 local research ethics committees. One response of these committees in general has been to establish executive subcommittees to deal in timely fashion with applications to multicentre research ethics committees. In Tully's experience this did shorten the time taken to respond, although less than a third of all local committees did so within the 21 days allowed, with a median time of 41 days.[3] After six months, Tully's study was still not approved by nine of the local committees. More worrying perhaps was that about half these committees asked for amendments, and two thirds of these concerned non-local issues—expressly against the Department of Health's guidance. Lux et al had a similar experience involving 99 local research ethics committees, with only a third replying within three weeks.[4] Some problems remained unsolved six months later. However, they did find, like Tully et al, that fast track subcommittees did speed up the process.[3] Al-Shahi and Warlow had a similar experience with a Scottish multicentre research ethics committee.[5] There the median delay to review was 28 days. The time taken for approval was 39 days, with a range of 21 to 109 days. They found only three objections, although one of these was not a local issue. The other major problem identified was the vast amount of paper involved—26.9 kg in one case[5] and over 100 000 sheets of paper in another.[3]

None of these studies looked critically at the workings of the multicentre research ethics committees themselves. There is one in each English region, and one each in Scotland and Wales. Their decisions are, however, binding for the whole of the United Kingdom. Any application involving five or more local research ethics committees goes first to the multicentre committee. So far, most of the problems seem to

BMJ 2000;320:1157–8