SYMPOSIUM ON EVIDENCE BASED MEDICINE

Research ethics and evidence based medicine

R K Lie

J Med Ethics 2004;**30**:122–125. doi: 10.1136/jme.2003.007203

In this paper, the author argues that the requirement to conduct randomised clinical trials to inform policy in cases where one wants to identify a cheaper alternative to known effective but expensive interventions raises an important ethical issue. This situation will eventually arise whenever there are resource constraints, and a policy decision has been made not to fund an intervention on cost effectiveness grounds. It has been thought that this is an issue only in extremely resource poor settings. This paper gives an example from the United Kingdom illustrating that this is also a problem faced by richer countries.

Correspondence to: R K Lie, Department of Clinical Bioethics, National Institutes of Health, Building 10, Room 1C118, Bethesda, MD 20852, USA; reidar.lie@fil.uib.no

Received 21 November 2003
Revised version received 21 November 2003
Accepted for publication 21 November 2003

It is increasingly becoming a standard requirement to document effectiveness before new interventions are adopted by a health care system. The methods of evidence based medicine, such as meta-analyses or other systematic evaluations of evidence, are often used to document effectiveness or the lack of it. There is an understandable reluctance to fund new experimental interventions when the evidence is not yet complete. Some of the most widely known controversies about the application of evidence based medicine, such as mammography screening for breast cancer, have been disputes over what the evidence says about effectiveness. Similarly, in discussions about the use of evidence based medicine for the issue of allocation of scarce resources, the focus has often been on the problematic assumptions made when using methods such as quality adjusted life years (QALYs).

In this paper I will argue that there is an additional issue which has not been addressed at all in the literature. In my discussion, I will assume that there is no disagreement about the established effectiveness of the intervention. I will further assume that there *is* agreement that the established effective intervention is too expensive for general implementation—that is, it is not cost effective. I recognise that both of these claims may more often than not be controversial, but for the sake of argument I will assume that there is no such controversy. The problem I want to address is how to establish the effectiveness of a *cheaper* intervention that can be widely implemented. According to evidence based medicine, we should conduct a randomised clinical trial with the proposed new, cheaper intervention in one of the treatment arms. I will show that this requirement raises an interesting and difficult ethical issue.

Perhaps the most famous recent example of an intervention that was considered too expensive to implement is the treatment of mother and baby to prevent perinatal transmission of HIV. An effective but very expensive intervention was established, which was unaffordable for developing countries. A proposed shorter, and cheaper, intervention was subsequently tested in randomised clinical trials in various countries, most of which included placebo in the control group. These trials were widely criticised because they did not use the established, proven treatment as a control. Those who defended the trials argued that the design was necessary to answer questions of relevance to developing countries. What is interesting about this discussion is that it has only been regarded as a problem for developing countries: in the developed world it has simply been taken for granted that one should use the established, effective treatment in the control group. In this paper I will argue that this is not true: the problem of the choice of control groups is a general one, arising in all situations of resource constraints, not just in situations of extreme lack of resources, such as in developing countries. It follows from this—because no country has unlimited resources for health care—that it is a general problem facing all of us. I will defend this claim by examining a recent case from the UK, involving treatment of multiple sclerosis.

NATIONAL INSTITUTE FOR CLINICAL EXCELLENCE AND TREATMENT WITH β INTERFERON FOR MULTIPLE SCLEROSIS IN THE UK

The National Institute for Clinical Excellence (NICE) in the UK was established with the aim of advising health professionals about which treatments work and which are cost effective. Although it has an advisory function only, its recommendations are quite influential. One important guiding idea behind the establishment of the Institute is the realisation that a government cannot afford to provide all expensive new treatments to everyone. The process by which particular new treatments are adopted have until now depended on accidental factors, such as the ability of patient groups and others to pressure politicians to accept funding for their disease. This can very easily lead to inequities, in that weaker patient groups do not receive the same consideration as patient groups with more powerful allies. Similarly, if reimbursement decisions are delegated to local authorities, one

Abbreviations: NHS, National Health Service; NICE, National Institute for Clinical Excellence; QALY, quality adjusted life year; MS, multiple sclerosis

can expect great variation in reimbursement practices. The evaluation process by NICE is supposed to ensure that only cost effective interventions are adopted by the National Health Service (NHS), and to ensure an equitable adoption of new interventions.

When NICE evaluates a new intervention, it will request relevant information from the manufacturer/sponsor of the technology or drug, and commission technological assessments from appropriate academic centres. The decision to evaluate a particular technology ultimately rests with the Ministry of Health. The commissioned reports are then reviewed by a committee set up by NICE, comments are solicited from interested parties, then reviewed again by the committee, and the final recommendation submitted.[1]

β interferon was licensed in the UK for treatment of relapsing multiple sclerosis (MS) in 1995. In order for a treatment to be reimbursed, a neurologist would have to certify that the drug would be clinically useful. In spite of this, many health authorities did not fund treatment with β interferon on grounds of unproven benefit and low cost effectiveness, leading to a charge of rationing by postcode. A High Court also ruled in favour of a patient who had obtained the necessary evaluation by a neurologist, arguing that ''a blanket ban was the very antithesis of a national policy, whose aim was to target the drug at patients who could most benefit from treatment''. This decision was similar to another High Court ruling with regard to Viagra—on a denial of reimbursement on the grounds of cost effectiveness. The Department of Health refused funding of the licensed drug for cost reasons. In 1999 a High Court ruled that this is against the law because it ''deterred doctors from exercising their duty to use their clinical judgment''. The Department of Health's policy was also deemed unlawful under European law because it ''contravened the so-called transparency directive on medicines, which lays down the principle that any decision to blacklist a medicine from a member state's national health service must state reasons based upon objective and verifiable criteria''. Subsequent draft guidelines then stated that Viagra could be prescribed for certain conditions, which presumably would fulfil the requirement of denial of treatment based on ''objective and verifiable criteria''. This constitutes the background both to the establishment of NICE and the particular appraisal of β interferon treatment.[1] NICE issued its preliminary appraisal on β interferon treatment for MS in October 2001, based on an evaluation process as described above. NICE did not recommend treatment with β interferon or glatiramer acetate for patients with MS. There was then a short appeal process, where an appeal board considered objections, but upheld the preliminary appraisal in January 2002.[2–4]

In the final appraisal, NICE did not recommend treatment of MS with β interferon or glatiramer acetate on the balance of the clinical effects and cost effectiveness. NICE recognised that scientific evidence shows that this treatment modality does reduce relapse frequency and severity (at least during the initial years of therapy) and is perceived to be of great value to people with MS, as the following quotation from the appraisal shows: ''The Committee considered in detail evidence taken directly from patients and two advocacy organizations. The patient organizations and the patients who attended the Committee meeting spoke of the patients' experience of this distressing disease and the impact of beta interferons and glatiramer on relapses and disease progression. This dialogue provided important insight into the effect of relapses on patients' daily lives and the value that they place on the potential reduction in severity of relapses with the use of these drugs.''

The Committee nevertheless concluded that the cost of the product is too high to justify its adoption, when one has to decide to use limited resources for competing interventions. Based on extensive economic modelling, the committee concluded that the cost per QALY gained would be between £248 000 and £810 000, for five years of treatment, and between £40 000 and £90 000 for 20 years of treatment.

The NICE appraisal process had several notable features:

- There was an initial cost effectiveness analysis. This showed great variations in cost effectiveness estimates, and the modelling was criticised. NICE then commissioned a more thorough cost effectiveness analysis.
- It is clear that the estimate of effects in relation to costs has a decisive influence on the recommendation that this intervention is not cost effective. The benefits estimates are done in terms of QALYs. The committee argues strongly that, although one may disagree with a number of things with regard to this methodology, it is the best summary measure of the benefits of an intervention, and those who criticise the way benefits are assessed should point to specific alternative evaluations.
- It is clear that there are major uncertainties in the data. There are only clinical trial data for about two years of treatment, and beyond five years there are no data at all. One therefore has to make estimates concerning treatment effects. Again, one might disagree with the specific choices that the committee has made, but there is no question that they have made a reasonable evaluation of the available evidence.
- It is also clear that whatever reasonable figures are used for the effects and costs of this treatment, it is going to be a very expensive one regarding benefits when compared with other funded treatments within the NHS.
- The process has been very public, with involvement of industry and patient advocacy organisations at every stage of the process. Key documents have been made available to all interested parties throughout the process.

In spite of this, complaints have been made from both sides about lack of transparency. Patient advocacy organisations have argued that they should have been present during deliberations. Others have complained that NICE is not able to make appropriate cost benefit calculations because it does not have access to key data that belong to the pharmaceutical companies, and that may be detrimental to their desire to have these drugs funded by the NHS.[5–7]

This, then, is an example of a recommendation not to adopt a treatment of proven benefit based on cost effectiveness data, and in spite of input and pressure from patient organisations to adopt the treatment. We may disagree with the specific conclusion, but in principle at least this is an example of the type of decision that my argument is based on: a decision not implement a treatment—not because it is of doubtful effectiveness, and not because one might disagree with some of the value judgments made, but because it is found to be too expensive given the limited resources available.

On 4 February 2002, however, the Department of Health announced that it would fund treatment with these two drugs, as part of a collaborative arrangement with the producers of the drugs to estimate long term effects. This development after NICE's decision is not important for the problem discussed here, but the ensuing discussion did provide some additional information about the basis for the original decision. Under the proposed scheme, all eligible patients as certified by a neurologist will receive reimbursement of the treatment, initially paid by the NHS. They will have to agree to be part of a monitoring programme of the drugs' effectiveness. The results will be fed into a statistical model, which is the same as the one used by NICE in its

appraisal, to calculate whether the results are better or worse than those used by NICE in its appraisal. If the results indicate that the cost effectiveness falls below a sum of £36 000 per QALY, industry will pay a proportional higher cost of the treatment. According to the proposal:[8] ''If actual benefit is equal to or greater than expected benefit (within a tolerance margin—see below) then the NHS will continue to make payment at the price agreed at the outset of the scheme. However if actual benefit after *t* years was below the tolerance margin, the price for the period up to the next review point will be reduced to the extent needed to restore cost effectiveness to the cost per QALY ''threshold'' determining entry into the scheme…, i.e. £36 000. We envisage that the formal monitoring process for assessing cost effectiveness and pricing adjustments will continue for up to 10 years. At the end of this period payments to companies will continue at the level implied by the final review point.''

The figure of £36 000 was arrived at in the following way: ''A retrospective analysis of appraisal determinations in its first year of operation, as summarised by Sir Michael Rawlins at NICE's annual public meeting, suggests that positive recommendations were in general associated with a cost per QALY of £30 000 or less; higher cost per QALY figures were accepted only if there were special factors accepted as relevant by the Appraisal Committee and not covered by the formal modelling. A number of ''special factors'' which might be considered to be relevant to the cost effectiveness of treatments for MS have been put to us in discussion. The FAD has specifically referred to two unquantified factors: (i) the impact of treatment on the severity (independent of the frequency) of relapses, and (ii) possible cost offsets from the avoidance of severe levels of disability requiring intervention by the Personal Social Services. In the light of all these considerations the threshold will be set, for the purpose of this scheme only, at £36 000.''

Patients immediately claimed victory, and saw this as a reversal of NICE's recommendation, although the Department of Health pointed out that there was no contradiction between this policy and the appraisal by NICE. NICE had urged the government to consider ways of making the treatments more cost effective, by, for example, negotiating price reductions, and thereby achieving an acceptable cost effective level. The proposed scheme would only mean that the NHS would pay the full amount for the treatment if it was found to be cost effective according to the set threshold. If the monitoring process should find that the benefits are less than that, industry will have to pay for part of the expenses of the drugs. The initial cost to the NHS in this scheme is estimated to be £7000–£10 000 per patient, with a total estimated annual cost of £50 million.

TESTING NEW POTENTIALLY COST EFFECTIVE INTERVENTIONS IN CLINICAL TRIALS

The fact that a government body, after a public process, whatever its flaws, has rejected a recognised clinically effective treatment on cost effectiveness grounds raises some interesting issues of research ethics which have not been explored fully. If an intervention is rejected on these grounds, it raises the question about what the appropriate control group should be in a future clinical trial, with the aim of identifying an effective but cheaper intervention for the condition.

Fundamental to research ethics is the concept of clinical equipoise. When testing a new, promising treatment in a randomised clinical trial, there should be no evidence that one of the treatments offered in the trial is more effective than the other. Associated with this idea is the claim, most famously connected with the debate about the revision of the Declaration of Helsinki, that the control group in a clinical trial is entitled to the best proven intervention. If there is clinical equipoise (the control group receives the best proven intervention, and the experimental group a new, promising treatment) it is felt that nobody who enters a clinical trial is disadvantaged. They will either receive an intervention they would ordinarily receive if they are randomised to the control group, or an intervention that, according to the best available evidence, is indistinguishable from the best proven treatment in terms of effectiveness. Therefore it would not, ordinarily, be appropriate to have a placebo group, or no intervention group, in a clinical trial where proven interventions are known to exist.

During the past few years there has been an extensive discussion about the choice of a control group in the context of research in resource poor settings. The question is whether it is permissible to use as a control group a treatment that is known to be less effective than the best proven therapy, in order to obtain knowledge that will be useful for the country in which the trial takes place. Thus, in the perinatal HIV transmission trials it was argued that it was permissible to have a placebo group to establish whether a short course treatment would be better than placebo—even though a known, effective treatment was already available in resource rich countries. This position has been highly controversial, as witnessed by the revision process of the Helsinki Declaration. The World Medical Association affirmed that the control group in a clinical trial should receive the best proven treatment, irrespective of where the trial takes place.

The discussion about the permissibility of choosing an intervention than is less effective than the best proven treatment, or best current treatment, has usually been seen as only relevant to resource poor settings. However, by reflecting on the situation after NICE rejected β interferon for the treatment of MS, it can easily be shown that the issue is one that applies in all cases of resource constraints—that is, is applicable to all countries.

Let us assume that the relevant government authorities in the UK decided that, all things considered, it should not reimburse β interferon treatment for MS, either as ordinary care or in the context of the research project described above. Let us also assume, given the wide consultation process before the decision, that there was general consensus in the country that β interferon treatment should not be covered, in light of all the other alternative uses of funds within the NHS. Given the unsatisfactory current treatment options of MS, it would be important to develop new but cheaper interventions. The issue is how one should design a trial for a promising, cheaper intervention in the future.

In such a trial there would be two choices for the control group. One could either provide the ''best proven treatment''—that is, β interferon treatment—or placebo. In light of the discussion about the revision of the Declaration of Helsinki, one might want to decide that placebo use would be unjustified. After all, there is a ''best proven'' treatment, albeit very expensive, that one would seem to be obligated to provide to the participants in the trial. From a methodological point of view, there is something to be said in favour of this strategy. We have identified a treatment that is known to be effective against MS, and we would therefore presumably want to know whether the new, promising treatment is at least as good as the established, effective, but too expensive treatment. If that is really what we are interested in, an equivalence trial would be unproblematic. However, we might very well expect that the new, cheaper treatment is not going to be as effective as the expensive, proven treatment, but still expect it to be better than what is currently offered within the NHS. If it is also the case—and this is a crucial point—that the course of untreated MS is highly variable, so that we do not know what the relapse rate

Table 1 Hypothetical trial results

Placebo/no treatment	New treatment	β interferon treatment
10*	11	12
9†	11	12

Numerals denote quality of life years. * Expected value; † Actual value.

is in a particular group of patients, an equivalence trial would not provide us with useful results. Let me defend this claim by using results from a realistic, but hypothetical trial, the results of which are provided in table 1.

Let us assume that from the data available we expect untreated MS patients of a certain age to expect 10 QALYs if they are left untreated. Let us also assume that if we treat them with β interferon we can increase their quality of life so that they now can expect to improve their prospects to 12 QALYS, but at a cost of £160 000, giving us a cost per QALY of £80 000, way over the accepted limit of £36 000. A new, promising, cheaper treatment costing £50 000 is tested against β interferon in an equivalence trial, showing an expected QALY of 11, which is in between the QALYs gained by β interferon and no treatment, giving a cost per QALY of £50 000, still above the accepted limit of £36 000. Based on this equivalence trial, we would therefore reject the new, promising treatment as not cost effective.

However, it turns out—unknown to us because we did not include a placebo group in the hypothetical trial—that the expected, untreated QALY in the trial population is actually 9 not 10. The cost benefit of β interferon would therefore be £53 000 in this patient group, still above the accepted limit. The cost benefit of the new treatment would, however, be £25 000, which falls within the accepted limit, and the new treatment should be adopted.

The important point is that we could not have known this if we had not included a placebo group in the hypothetical trial. If we had done an equivalence trial we would have wrongly rejected the new treatment as not cost effective, whereas if we had included a placebo group, we would have seen that the new treatment is indeed better than the current acceptable treatment, which in this case is no treatment.

It is evident that the structure of this hypothetical trial is exactly the same as the structure of the perinatal HIV transmission trials, showing that problem of a choice of a control group in the presence of established effective interventions is the same in all settings of resource constraints, not just in settings of extreme scarcity of resources.

In this case, since there has been a decision not to introduce the intervention in the National Health Service, one might want to argue that one is not denying anybody anything they have a claim to, in spite of the requirement of the Declaration of Helsinki. In the UK, however, the health authorities decided to introduce β interferon treatment for MS in the context of a research project, in an attempt to gather data about the long term effects of β interferon treatment, as well as limit the costs to the NHS should the treatment turn out to be quite cost ineffective. As some commentators have noted, the value of the research project is quite doubtful as it is going to be difficult to establish reliable effectiveness data in the absence of reliable comparative data about the course of MS without treatment. Quite apart from that problem, there is an additional problem if the scheme should result in a judgment that β interferon treatment does indeed fall outside the accepted limit of cost effectiveness. One would then also have to search for a cheaper, but still effective treatment, and the issue of the choice of a control group would return. However in this scenario, all eligible patients would then be receiving the expensive treatment, and one would presumably have to take some off this treatment in order to establish the cost effectiveness of the new, promising treatment. This is clearly morally more problematic that not providing something to the control group that they would not have received anyway from the national system.

CONCLUDING REMARKS

The case of β interferon treatment for multiple sclerosis shows that we face the problem of the choice of an appropriate control group whenever there are resource constraints and whenever equivalence trials cannot be carried out for scientific reasons. The debate about what is an appropriate control group is therefore not only relevant to developing countries, but is going to be increasingly important for all countries as pressure increases to prioritise among expensive, new interventions. This problem raises the issue of the relativity of ethical standards in research. If we assume that all countries face exactly the same resource constraints, the design of trials will be the same everywhere. However, if we assume that resource constraints will vary significantly, even in relatively resource rich settings, treatment access will vary in different countries, necessitating different standards of care. Although this might be considered ethical relativism by some, it should more appropriately be seen as adapting universal ethical principles to different local circumstances. What the β interferon case shows is that it is likely that local circumstances may vary more dramatically than we have been used to so far, as we are facing increasingly expensive, but effective interventions.

The views in this paper are the author's and do not represent the policies or positions of the National Institutes of Health, the Public Health Service, or the Department of Health and Human Services.

REFERENCES

1 **Dillon A**, Gibbs TG, Riley T, *et al*. The National Institute for Clinical Excellence and coverage for Relenza by the NHS. Informing judgment: Case studies of health policy and research in six countries, Milbank Memorial Fund, 1999.
2 **National Institute for Clinical Excellence**. Beta interferon and glatiramer acetate for the treatment of multiple sclerosis. London: National Institute for Clinical Excellence, 2002.
3 **National Institute for Clinical Excellence**. Appraisal of the use of beta interferons and glatiramer acetate in the treatment of multiple sclerosis. London: National Institute for Clinical Excellence, 2002.
4 **National Institute for Clinical Excellence**. NICE appraisal of beta interferon and glatiramer for MS. Process to 2002. London: National Institute for Clinical Excellence, 2002.
5 **Cookson R**, McDaid D, Maynard A. Wrong sign, nice mess: Is national guidance distorting allocation of resources. *BMJ* 2001;**323**:743–5.
6 **Smith R**. The failings of NICE. Time to start to work on version 2. *BMJ* 2000;**321**:1363–4.
7 **Sculpher M**, Drummond M, O'Brien B. Effectiveness, efficiency and NICE. A nice start but evidence costs money. *BMJ* 2001;**322**:943–4.
8 **Department of Health**. Cost effective provision of disease modifying therapies for people with multiple sclerosis. London: Department of Health, 2002.