Preferences and understanding their effects on health*

K McPherson, A Britton

Abstract

Preference for a particular intervention may, possibly via complicated pathways, itself confer an outcome advantage which will be subsumed in unblind randomised trials as part of the measured effectiveness of the intervention. Where more attractive interventions are compared with less attractive ones, any difference could therefore be a consequence of attractiveness and not its intrinsic worth. For health promotion interventions this is clearly important, but we cannot tell how important it is for therapeutic interventions without special studies to measure or refute such effects. These are difficult to do and are complex. Until the therapeutic effects of preference itself are more clearly understood, understanding the true therapeutic effects will be compromised, at least in principle.

(*Quality in Health Care* 2001;**10**(Suppl I):i61–i66)

Keywords: patient preference; randomised controlled trials; measurement

Key messages

- Preferences for health interventions may affect outcomes.
- Randomisation between important preferences is not possible, so reliable measurement of preference effects is difficult.
- Where preferences exist, they may contaminate unblinded randomised controlled trials.
- Randomised comparisons are usually thought to estimate therapeutic benefit.
- This requires knowledge about the lack of differential preference effects, and such information is generally lacking.
- Unblind trials are therefore not necessarily properly controlled.

It is self-evident that public health policy, healthcare strategies, and medical treatments are best if based on hard evidence of attributable effect. When assessing these effects in research studies we want to be able to measure accurately how much of any observed health improvement is due to the intervention itself. However, this is known to be fraught with problems about the nature of evidence and the methodologies used.[1 2]

One of the main methodological issues in research design is controlling for confounding variables that are linked both to the intervention and to the observed outcome—for example, age, prognosis, and social class. Randomised controlled trials (RCTs) are usually regarded as the "gold standard" design for obtaining hard evidence for a therapeutic or behavioural effect on outcome.[3] Participants are randomised between intervention groups so that any known and unknown confounding factors will be equally distributed between the groups, and any outcome differences observed between the groups can be attributed with more certainty to the difference between the interventions. In order to reduce other biases of preferences or measurement the trial participants and the researchers are rendered unaware of which intervention or treatment is being given. A "double blind" trial is relatively easy to conduct when the active treatment is a drug for which an inactive but visually identical placebo can be prepared. It is much less easy in trials of health promotion interventions and many surgical procedures.[4]

Non-randomised studies (observational studies), in which participants are typically allocated to interventions according to the opinion or preference of their practitioner and their own preferences, need to be interpreted with appropriate caution. Treatment choices can depend very much on individual characteristics that may include prognosis, and discerning the difference between a real effect of treatment or an effect of confounding by selection is normally very difficult. Many examples exist in which non-randomised studies indicate a substantial "treatment" benefit that is subsequently refuted once all confounding has been eliminated by randomisation.[5]

In this paper we will argue that, just as non-randomised studies may provide unreliable evidence of effect unless very strict conditions are met,[6] so RCTs that are not "blinded" are also difficult to interpret. There are many large RCTs evaluating new treatments in cancer or heart disease that are, of necessity, unblinded and hence cannot—by the arguments to follow—provide reliable evidence for a treatment effect as we understand it. People may have strong preferences for different treatment options and for health enhancing interventions such as diet, exercise, and tobacco. In these circumstances, where these preferences may influence the effectiveness of interventions through psychological or immunopsychological pathways or through compliance with the intervention, unblind trials may wrongly attribute effects solely to the physiological or pharmacological properties of an intervention.

We will argue that, unless and until the nature, extent, and size of such preference effects on the effectiveness of interventions is better understood, unblinded randomised

Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London WC1 7HT, UK
K McPherson, *professor of public health epidemiology*

Department of Epidemiology and Public Health, University College London, London WC1E 6BT, UK
A Britton, *lecturer*

Correspondence to:
Professor K McPherson
k.mcpherson@lshtm.ac.uk

*This paper is an amended version of Chapter 9 in *Health promotion—new discipline or multi discipline* (Edmondson R, Kelleher C, eds), published by the Irish Academic Press, Dublin, 2000.

Qual Health Care: first published as 10.1136/qhc.0100061.. on 1 September 2001. Downloaded from http://qualitysafety.bmj.com/ on March 28, 2024 by guest. Protected by copyright.

studies do not necessarily provide secure information about the effects of interventions themselves. This is not a critique of randomisation but points to a neglected, but possibly crucial, component of the accurate interpretation of evidence.

Patient preferences

When we say that an intervention "works", we are usually referring to an average beneficial effect among a group. The benefit in an individual or in a subgroup might be different, but such complexities of *interaction* are generally neglected in attempts to discern the overall main effects. One particular kind of interaction of interest here is that between a biological response to a treatment or intervention and a possible psychological or social effect of receiving that treatment/intervention. An example of this particular interaction might be the therapeutic effect of patient preferences,[7] part of the clearly related phenomenon of the placebo effect.[8]

Interventions work in clinical medicine because of specific effects on a biological mechanism. Thus, drugs which kill rapidly dividing cells or compounds which have antihormonal effects can be expected to go to the root of a particular pathological process which will be common to all who suffer from it. Public health interventions tend to be less specific. An intervention designed to enable people to improve their health by, for example, a change in lifestyle does not have the same kind of simple (in principle) underpinning and theoretical justification because the complex process of making a change interacts with the health consequence of that change, in addition to the complex social and contextual determinants of behaviour in the first place. The health effects may well have a simple biological underpinning, but the roots of the process of behavioural change are much less likely to do so.

Believing in the effectiveness of a treatment or having a preference for one procedure over another may influence the biological effect by several plausible mechanisms. An individual, if randomised to his or her preferred intervention, may experience an enhanced response above and beyond the average biological effect of the intervention. If so, this will complicate rigorous interpretation of the effectiveness of the intervention. However, it is unavoidable that some people will have preferences for one intervention or another and recruitment to randomised trials may therefore become problematic.[9 10] Even if people agree to randomisation, the ability to detect preference effects is intrinsically compromised.[11] Firstly, one can never randomise between enthusiasm for a treatment and hating the whole idea of it. In other words, randomisation between preferences that matter is never possible, which is why randomisation is important—to attempt aggregation over preferences, to detect the main treatment effect.

To compare genuine preferences, the possibility of confounding must be considered; people who prefer something may be different in other ways, plausibly related to the prognosis, from those who do not.[12] Because of these difficulties, examination of preference effects is neglected. However, it is vital to disentangle the therapeutic effect on health enhancement from any possible benefit of preferences themselves.[13] Randomised trials, whose main purpose is to detect aggregate biological effects, start by having to find a way of overcoming or circumventing unmeasured preference effects otherwise they cannot detect the therapeutic effects reliably. This is why blinding is important. Unblind randomisation is not in itself sufficient to eliminate these effects if they are real.

Evidence for preference effects

In health promotion it is obvious that the most beneficial changes will occur mainly in those most willing and able to change in response to an intervention that accords with their context, since empowering people in their own context is effective health promotion.[14] However, there is little irrefutable evidence for significant psychological preference effects in medical interventions on health outcome.

A well known example, the Coronary Drug Project,[15] may provide some clues. Treatment for the secondary prevention of coronary heart disease in middle aged men was observed to be dramatically more effective at delaying death when a placebo was "properly" taken than when not. If drug compliance here is a measure of some enthusiasm for the intervention, and non-compliance a measure of little enthusiasm, then individual preferences may seem to have an important effect on outcome which is not strictly or directly pharmacological. Adjusting for 40 potential confounding variables, most of which were physiological measurements, when comparing adherers with non-adherers made little difference to the comparison. A preference effect rather than confounding is easily the most plausible explanation. Since the drug was a placebo, by definition the therapeutic effect was zero, but the adjusted difference in the 5 year mortality between adherers and non-adherers was 16% and 26% which was highly significant.

Although this example compares adherence with intervention, it is not plausibly adherence itself that is responsible for the lower mortality since the drug (placebo) was designed to be inactive, whether taken or not. It remains possible that adherence predicts compliance with other incidental aspects of care, but it is difficult to imagine what might have such a large effect. This example provides some evidence that there might be an important effect of belief in intervention and consequent lower mortality, which could be large.

Another example comes from randomised trials of drug treatment for postpartum pain in Norway. Two trials were run consecutively in which the first compared active drug with placebo and the patients were told that they may get the placebo, and the second compared two active drugs. In the second trial the drug common to both trials had a greater response, in terms of less pain intensity, when patients were aware that no placebo was involved.[16]

Qual Health Care: first published as 10.1136/qhc.0100061.. on 1 September 2001. Downloaded from http://qualitysafety.bmj.com/ on March 28, 2024 by guest. Protected by copyright.

Table 1 Additive model of preferences

Postulated intervention effects if:	*Indifferent*	*Prefer A*	*Prefer B*
On intervention A	*P*	*P+y*	*P–y*
On intervention *B*	*P+x*	*P+x–y*	*P+x+y*

There is a great deal of other indirect evidence for such effects, none of which is conclusive, and an equally fascinating debate on the possible biological pathways for such psychogenic effects.[17–20] If such effects exist, we must first understand the consequences and then try to understand when such effects might be important. However, since randomisation between preferences is not possible, evidence has to be observational and hence dogged by possible (indeed, highly likely) unmeasured confounding. This is a complex and difficult scientific agenda—is it worth the effort?

A simple additive model

To examine the possible effect of preferences in straightforward unblind trials, a simple model will illustrate why these factors are important. Imagine two interventions designated by *A* and *B*. Assume that "biologically" or, strictly, organically, intervention *A* affects (that is, works on) a proportion *P* of eligible people, and intervention *B* affects a higher proportion *P+x*. This might be the cure rate, the proportion surviving for 5 years, or simply the proportion of people affected positively by the intervention in the absence of any effect of preferences. Clearly, the complexities in this simplification are profound, but we start somewhere recognisable.

Assume also that having a preference for (or being prepared to respond to) *A* bestows an extra average advantage for intervention *A* of an amount *y* to *P+y*. Here y is the unknown effect of preference. Similarly, a preference for *B* bestows a similar additive effect *y* to *P+x+y* for intervention *B*. Conversely, of those who prefer *A*, only *P+x–y* will be affected if given intervention *B*, and of those who prefer *B*, *P–y* will benefit if given *A*. Of the people who are indifferent to the interventions, *P* is the proportion who benefit on *A* and *P+x* benefit on *B*. These are postulated average effects of that group of patients for whom these interventions would be appropriate, obviously with an interaction of preferences. These effects are summarised in table 1.

If the proportion of the eligible population who prefer intervention *A* is α, if β prefer *B*, and γ are indifferent, we will require that (α+β+γ) =1. More complicated models could be imagined in which the effects of preference were graded, multiplicative, or asymmetrical, but since the effects of preference are poorly understood, the simplest possible model is to be preferred.

It can be shown that, in a large well conducted randomised comparison, on the above assumptions the unbiased *estimate* of the effect of intervention *B* over intervention *A* (by subtracting the mean effect in those randomised to *B* from that in those given *A*) will be:

$$x + 2y\,(\beta - \alpha)$$

This will be different from *x* (the true biological/physiological effect) by an amount equal to $2y\,(\beta - \alpha)$ (the preference component). Hence, in this simplest model, such an RCT will only estimate *x* correctly either if $y = 0$ (no preference effects) or if $\beta = \alpha$ (that is, the proportions favouring *A* and *B* are the same).

The question then is: by how much will the effect of an intervention be overestimated or underestimated as a result of preference effects under reasonable assumptions on *y* and on $(\beta - \alpha)$, and what might this do to our understanding of evidence based health care and health promotion? Firstly, we have to postulate the size of the difference in the proportions preferring the two interventions. If 5% prefer intervention *B* and 30% intervention *A*, the difference is 25%, that is $(\beta - \alpha) = -0.25$. Secondly, we need to estimate the additional biological advantage of having treatment *B* over treatment *A* (that is, *x*). For illustrative purposes, let us imagine this to be 10% and the effect of treatment *A* alone is to benefit 50%. If the preference advantage (*y*) is 5% (as it probably was in the Coronary Drug Project, $2y$ = 26% – 16%, above), then the intervention estimates of effect are as shown in table 2. If these values are ever appropriate, then simple substitution will indicate that a fair RCT will be 25% "out" and, in this case, $2y\,(\beta - \alpha)$ = 25% of *x*—that is, *x* (the "intervention" effect) will be estimated as $0.75x$. If, on the other hand, 30% prefer *B* and 5% *A*, then the RCT estimate will be $1.25x$. Either way, both these results are wrong in the sense in which they would be understood by a conventional interpretation of such a trial, ignorant (as we must usually be) of the size of the effect of patient preferences. These estimated differences will be taken as attributable to the intervention alone, and hence generalisable to other groups of people.

Clearly, if the difference in the proportions who prefer *A* or *B* rises to 50%, the size of the "bias" from a randomised comparison rises itself to 50% for these hypothetical values of x=10% and y=5%. If, however, *y* is only 1% (a smaller preference effect), then the "biases" in the results of the RCTs will be reduced to 5% for a 25% difference in proportions with contrasting preferences, and 10% for a 50% difference. However, if *y* is 10%—that is, the role of preference is more profound than the intrinsic intervention effect—then the trials will be, respectively, 50% and 100% "out" on average—that is, the intervention effect will be estimated as $1.5x$ or $2x$). This is clearly non-trivial if such large differences are plausible in the prevalence of important (i.e. potent) preferences in trials where blinding is not possible or deemed unnecessary.

Thus, if preference effects are real but evenly balanced, they may not distort randomised comparisons by much. But, unlike in strictly biological mechanisms, preferences may be

Table 2 Intervention effects of model

Postulated effects of intervention if:	*Indifferent*	*Prefer A*	*Prefer B*
On intervention A	50%	55%	45%
On intervention B	60%	55%	65%

Qual Health Care: first published as 10.1136/qhc.0100061.. on 1 September 2001. Downloaded from http://qualitysafety.bmj.com/ on March 28, 2024 by guest. Protected by copyright.

transient and may change and thus may not be evenly balanced all the time. The consequence of these effects will be attenuated by the amount of balance exhibited in a single trial, and often even the relative attractiveness of interventions will be poorly understood.

In health promotion a much discussed theory considers stages of change[21] in which a progression is described in people ranging from no inclination whatever to alter behaviour (precontemplation) to being entirely prepared to change behaviour (contemplation or preparation). The proportion in the latter category may strongly predict or, indeed, determine the apparent effectiveness of interventions, quite apart from their intrinsic merits in aggregate. The net outcome of the RCT will therefore be affected by the proportion prepared to change behaviour as well as the differential effectiveness of the two interventions among the different categories of preparedness. In a randomised comparison the proportions will not be stochastically similar because different interventions might have very different categories of people in contemplation to change behaviour than for competing interventions. This is probably qualitatively different from therapeutic medicine.

This phenomenon will necessarily give rise to confusion about the true mechanisms involved, when one has no idea a priori for which categories such interventions work and what proportion of any population exhibits these categories of behaviour. It might also give rise to systematic and important confusion about the nature and extent of such effects. What are the possible effects of strictly adhering to randomised comparisons as the gold standard to estimate the true effectiveness of interventions, when such a strategy clearly ignores potent choices and preferences in order to identify a mean therapeutic effect? How large a bias can be induced by such a strategy, and how much confusion?[22]

We learn several things from this simple analysis, although reality in health care is likely to be much more complex than this. From this simple example it is clear that preferences may distort unblind randomised comparisons where they do have a therapeutic effect. More importantly, they may sometimes be entirely responsible for the putative effects of treatments, which are actually equivalent in a strictly and direct therapeutic sense.

Conclusion

It is interesting to speculate on the implications of this on the dynamics of understanding the effectiveness of health care. An intervention that is thought to be beneficial may be beneficial largely because of belief in its merits. Moreover, that belief is likely to be reinforced by such an overestimation of the true "biological" effect in the absence of any coherent understanding of preference effects.

Enthusiastic doctors will be eager to help their patients with a chronic disease to get better. If a new intervention is on offer it will only be proposed if it confers some important theoretical benefit in the eyes of the advisor. Although Chalmers[23] has shown that new interventions are just as likely to be worse rather than better than their predecessors, such a finding is unlikely to be widely absorbed either by enthusiastic clinicians or by their patients who want to get better. We might therefore assume that, in general, new interventions have much more enthusiastic support than old interventions (that are clearly not magic bullets) in chronic diseases such as cancer and heart disease that are not always amenable to intervention.

Clearly, belief in an intervention does not necessarily enhance its effectiveness, but it might and, if it does, we would in general be very unlikely to know. On these two assumptions—that belief does affect effectiveness and that people faced with poor therapeutic prospects are likely to believe in an untried intervention more than a conventional one—organically ineffective interventions are, on average, likely to gain in apparent effectiveness as "unbiased" evidence accumulates from RCTs and more people prefer them. If so, it is important that this process does not simply and systematically accrue more and more expensive (and possibly nasty) placebos, unless we are sure there is no other way to enhance the effect of care.

So long as the preference component of effect ($2y$ above) is smaller than the biological/physiological component, giving people interventions they do not prefer may still confer some important advantage. However, whenever the preference component is large, trials can badly misrepresent the "true" effect and favouring one intervention because of its apparent aggregate effect may cause some to suffer a net disadvantage.

Many things follow from this analysis and it behoves us to understand better the nature of the phenomenon. We need to know where preferences are important and where they are not, and what the implications of this are on our understanding of the results of unblind randomised comparisons. For example, interventions designed to affect behaviour as the main outcome might well be much more susceptible to important preference effects than those with organic outcomes such as death. In general, preferences are likely to be very different for some interventions than for others, and this may be complicated by possible class, sex, age, geographical or ethnic variations.

Just as the interpretation of non-randomised studies is problematic unless confounding by prognosis is known to be absent, the interpretation of unblind RCTs requires a further caution. Since they represent a neat solution to the former concern, we need either to develop another methodological panacea for dealing with preference effects or to understand them reliably.

What to do next

The obvious progression is to mount randomised trials from which the size of preference effects can be reliably measured. Rucker[24] has postulated a two stage design where

Qual Health Care: first published as 10.1136/qhc.0100061.. on 1 September 2001. Downloaded from http://qualitysafety.bmj.com/ on March 28, 2024 by guest. Protected by copyright.

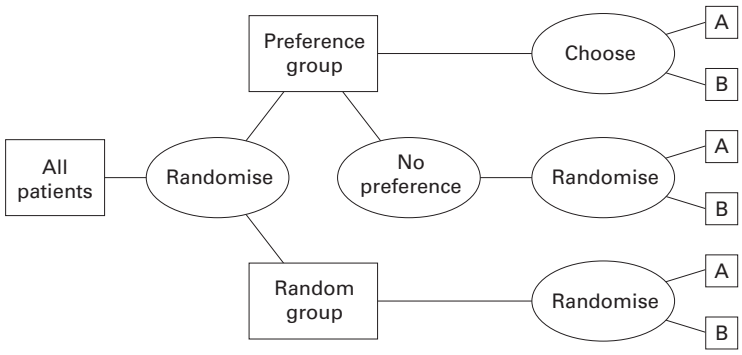


Figure 1 Rucker design to determine preference effect.

randomisation is between two groups—one a random allocation between interventions and the other a choice of intervention randomised between those with a preference and those without (fig 1). Thus, the two arms compare no choices with patient preferences where they exist. The problem here is a practical one of interpretation as subtraction of the means from the two randomised groups provides an estimate of a complex algebraic function of the main physiological effects x and any preference effect $2y$.

The physiological effect itself will have to be estimated from the randomised arm, but with an unknown preference component, as above, based on imprecise estimates of the proportions α and β from the preference arm. A preference effect might just be estimable from comparison of the results from the two arms. The algebra, which is laborious but simple (and available from the authors), convinces one again that estimating these effects reliably is difficult and would require large numbers of patients for sufficient precision.

While perhaps easier to put into practice and ethically sound, the trial design described by Brewin and Bradley[25] will produce results for which a preference effect cannot be disentangled from the possible confounding arising from systematic differences between patients with strong intervention preferences and those without (fig 2). This puts the inference back to observational (non-randomised) studies, and hence no methodological progress is made.

Torgerson and colleagues[26] argue that determining preference effects can be achieved by recording preferences in a randomised trial—that is, by asking participants which treatment they would prefer if they were given the choice—and fitting a preference term as an interaction in a regression analysis of main treatment effects. This will, however, be insensitive if the numbers in the trial are determined solely by the necessity to detect a main effect of a particular size, since interactions usually require an order of magnitude more patients for the same power. Thus, if trials more than 10 times the size necessary to detect a main effect can be ethically justified, then detecting preference effects may be feasible. Attempts to detect such an effect by this means have either been underpowered or, in a study of exercise for back pain, failed to find any effect.[27] Another attempt, in this case including a non-randomised patient preference arm for counselling or cognitive behaviour therapy for depression, failed because the selection processes for treatment of depression are poorly understood.[28]

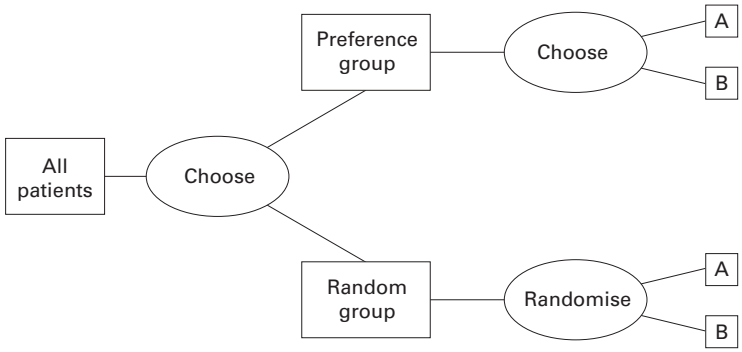


Figure 2 Brewin and Bradley preference design.

Large trials usually require a convincing biological or clinical justification to obtain enthusiastic support, where the "biomedical model of disease is so pervasive that we often fail to see it as such, but view it as a reality".[29] Invoking alternative mechanisms that sound slightly fanciful themselves and for which there is not much evidence are therefore unlikely to be enough on their own. It remains to be seen whether sufficient evidence can be derived from what we already know to justify preference effect trials in any area. If it cannot, the much larger question will be whether it can ever be derived from what we can ever know and agree upon, however hard we try. Perhaps the positivist discourse has not so much to recommend it after all. But we certainly do need to know eventually! How much of health care or health promotion is "merely" a placebo is, after all, quite an important question. The recent systematic review from York suggests that doctors can induce better outcomes by their own warmth and reassurance,[30] but what is now required is a more people centred focus which measures their own effects on their own outcomes. These are likely to be more potent still.

1 Sackett DL, Wennberg JE. Choosing the best research design for each question. *BMJ* 1997;**315**:1636–7.
2 McPherson K. The best and the enemy of the good: randomised controlled trials, uncertainty, and assessing the role of patient preference in medical decision making. The Cochrane Lecture. *J Epidemiol Community Health* 1994;**48**: 6–15.
3 Kleijnen J, Gotzsche P, Kunz RA, *et al*. So what's so special about randomisation? In: Maynard A, Chalmers I, eds. *Non random reflections on health services research*. The Nuffield Provincial Hospitals Trust. London: BMJ Books, 1997: Chapter 4.
4 McKee M, Britton A, Black N, *et al*. Choosing between randomised and non randomised studies. In: Black N, Brazier J, Fitzpatrick R, Reeves B, eds. *Health services research methods*. London: BMJ Books, 1998: Chapter 6.
5 Omenn GS, Goodman GE, Thornquist MD, *et al*. Effects of a combination of beta carotene and vitamin A on lung cancer and cardiovascular disease. *N Engl J Med* 1996;**334**: 1150–5.
6 Sanderson C, McKee M, Britton AR, *et al*. Randomised and non-randomised studies: threats to internal and external validity. In: Stevens A, Abrams K, Brazier J, eds. *Advanced handbook of methods in evidence based healthcare*. London: Sage, 2001:95–115.
7 Antonovsky A. Can attitudes contribute to health? *Advan Mind-Body Health* 1992;**8**:33–48.
8 Kleijnen J, de Craen AJM, Everdingen JV, *et al*. Placebo effect in double-blind clinical trials: a review of interactions with medications. *Lancet* 1994;**344**:1347–9.
9 Lilford RJ. Equipoise and the ethics of randomisation. *J R Soc Med* 1995;**88**:552–9.

Qual Health Care: first published as 10.1136/qhc.0100061.. on 1 September 2001. Downloaded from http://qualitysafety.bmj.com/ on March 28, 2024 by guest. Protected by copyright.

10 MacIntyre IMC. Tribulations for clinical trials. Poor recruitment is hampering research. *BMJ* 1991;**302**:1099–100.
11 Silverman WA, Altman DG. Patient preferences and randomised trials. *Lancet* 1966;**347**:171–4.
12 McKay JR, Alterman AL, McLellan, *et al*. Effect of random versus nonrandom assignment in a comparison of inpatient and day hospital rehabilitation for male alcoholics. *J Consulting Clin Psychol* 1995;**63**:70–8.
13 McPherson K. Patients' preferences and randomised trials. *Lancet* 1996;**347**:1119.
14 World Health Organization. *Targets for health for all 1991–1993*. Health for All Series No. 1. Copenhagen: WHO Regional Office for Europe, 1985.
15 The Coronary Drug Project Team. Influence of adherence to treatment and response of cholesterol on mortality in the Coronary Drug Project. *N Engl J Med* 1980;**303**:1038–41.
16 Skovlund E. Should we tell trial patients that they might receive placebo? *Lancet* 1991;**337**:1041.
17 Phillips DP, Ruth TE, Wagner LM. Psychology and survival. *Lancet* 1993: **342**:1142–5.
18 Redd WH, Silverfarb PM, Anderson BL, *et al*. Physiologic and psychobehavioral research in oncology. *Cancer* 1991; **67**:813–22.
19 Cassileth BR, Lusk EJ, Miller DS, *et al*. Psychosocial correlates of survival in advanced malignant disease. *N Engl J Med* 1985;**312**:1551–5.
20 Levy SM, Herberman RB, Lee J, *et al*. Persistently low natural killer cell activity, age, and environmental stress as predictors of infectious morbidity. *Nat Immun Cell Growth Regul* 1991;**10**:289–307.
21 Sheppard BH, Hartwick J, Warshaw PR. The theory of reasoned action: a meta analysis of past research with recommendations for modifications and future research. *J Cons Res* 1988;**15**:325–39.
22 McPherson K, Britton AR, Wennberg J. Are randomised controlled trials controlled? Patient preferences and unblind trials. *J R Soc Med* 1997;**90**:652–6.
23 Chalmers I. What is the prior probability of a proposed new treatment being superior to established treatments? *BMJ* 1997;**311**:74–5.
24 Rucker G. A two-stage trial design for testing treatments, self-selection and treatment preference effects. *Stat Med* 1989;**8**:477–85.
25 Brewin CR, Bradley C. Patient preferences and randomised clinical trials. *BMJ* 1989;**299**:313–5.
26 Torgerson JT, Klaber-Moffett J, Russell IT. Patient preferences in randomised trials: threat of opportunity. *J Health Serv Res Policy* 1996;**1**:194–7.
27 Klaber-Moffett J, Torgenson D, Bell-Syer S, *et al*. Randomised controlled trial of exercise for low back pain: clinical outcomes, costs and preferences. *BMJ* 1999;**319**:279–83.
28 Ward E, King M, Lloyd M, *et al*. Randomised controlled trial of non-directive counselling, cognitive behaviour therapy, and usual general practitioner care for patients with depression. *BMJ* 2000;**321**:1383–8.
29 Phillips DP, Ruth TE, Wagner LM. Psychology and survival. *Lancet* 1993;**342**:1142–5.
30 Blasi, ZI, Harkness E, Ernst E, *et al*. Influence of context on health outcomes: a systematic review. *Lancet* 2001;**357**: 757–62.

Qual Health Care: first published as 10.1136/qhc.0100061.. on 1 September 2001. Downloaded from http://qualitysafety.bmj.com/ on March 28, 2024 by guest. Protected by copyright.