Lay conceptions of the ethical and scientific justifications for random allocation in clinical trials

Elizabeth J. Robinson[a,*], Cicely Kerr[a], Andrew Stevens[b], Richard Lilford[b], David Braunholtz[b], Sarah Edwards[c]

[a] *Department of Psychology, Keele University, Keele, Staffs ST5 5BG, UK*
[b] *Department of Public Health and Epidemiology, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK*
[c] *Centre for Ethics in Medicine, University of Bristol, Bristol BS2 8BH, UK*

Abstract

Randomised controlled trials (RCTs) play a central role in modern medical advance, and they require participants who understand and accept the procedures involved. Published evidence suggests that RCT participants often fail to understand that treatments are allocated at random and that clinicians are in equipoise about which treatment is best. We examine background assumptions that members of the public might draw upon if invited to take part in a RCT. Four studies ($N = 82$; 67; 67; 128), in the UK, identified whether members of the public (i) accept that an individual clinician might be genuinely unsure which of two treatments was better; (ii) judge that when there is uncertainty it is acceptable to suggest deciding at random; (iii) recognise scientific benefits of random allocation to treatment conditions in a trial. Around half the participants were loathe to accept that a clinician could be completely uncertain, and this was no different whether the context was one of individual treatment or research. Most participants found it unacceptable to suggest allocating treatment at random, though there was weak evidence that a research context may reduce the unacceptability. Participants did not judge that more certain knowledge would be gained about which treatment was best when treatments were allocated at random rather than by patient/doctor choice: scientific benefits of randomisation were apparently not recognised. Judgements were no different in non-medical contexts. Results suggest a large mismatch between the assumptions underlying the trial design, and the assumptions that lay participants can bring to bear when they try to make sense of descriptive information about randomisation and equipoise. Previous attempts to improve understanding by improving the clarity or salience of trial information, or of making explicit the research context, while helpful, may need to be supplemented with accessible explanations for random allocation.
© 2003 Elsevier Science Ltd. All rights reserved.

Keywords: UK; Clinical trials; Randomised control trials; Participation; Randomisation; Equipoise

Introduction

Randomised controlled trials (RCTs) are widely considered to be the best method of advancing knowledge about the effectiveness of medical treatments, and their incidence has increased from 2116 trials listed in 1960 to 348,740 in 2002 (Cochrane Controlled Trials Register, 2002). These trials rely on millions of patients' consent to participate. Consent or refusal to participate in such trials is adequately informed only if patients grasp two key features in addition to the details of what will happen to them in their particular trial: that participants will be allocated randomly to treatment arms, and that at the start of the trial there are no convincing grounds for supposing that any patient would be advantaged or disadvantaged if allocated to one treatment arm rather than another. This consensus state of uncertainty amongst clinicians has been labelled clinical equipoise (Freedman, 1987) or collective equipoise (Johnson, Lilford, & Brazier, 1991).

*Corresponding author.

E-mail address: e.j.robinson@keele.ac.uk (E.J. Robinson).

There is ample evidence in the published literature that trial participants often show signs of misunderstanding the basis of their treatment allocation, and of wrongly assuming that one treatment is already known to be better than the other or others. Many researchers find a relatively high incidence of trial participants' failure to report that their treatment was allocated at random (e.g. Glogowska, Roulstone, Enderby, Peters, & Campbell 2001; Hietanen, Aro, Holli, & Absetz, 2000; van Stuijvenberg et al., 1998; review by Edwards et al., 1998). Several studies report failures to acknowledge equipoise (e.g. Ellis & Butow, 1998; Edwards et al., 1998).

Importantly though, being able to report accurately that treatments were allocated at random does not necessarily indicate coherent understanding. Some participants state that treatments were allocated at random, but nevertheless reveal other inconsistent beliefs by reporting that they were assigned treatment according to their particular symptoms (Appelbaum, Grisso, Frank, O'Donnell, & Kupfer, 1999; Appelbaum, Roth, Lidz, Benson, & Winslade, 1987; Featherstone & Donovan, 1998, 2002; Snowdon, Garcia, & Elbourne, 1997). Some participants construct their own incorrect explanation for random allocation, such as that it removes responsibility from individual doctors or parents, or that it provides a way of rationing a scarce or costly resource (Snowdon et al., 1997). Both these kinds of misconception—that doctors assigned treatment according to individuals' symptoms, or that random allocation is used for non-scientific reasons— imply a lack of recognition of equipoise. The interviews in Appelbaum and colleagues' studies (1987, 1999) took place immediately after trial information was given. By contrast, interviews took place up to 2 years after consent in the case of Snowdon et al. (1997). Featherstone and Donovan (1998, 2002) interviewed some participants after 6 months, after treatment was complete. Although the anger and upset which some of these participants revealed in their interviews do give cause for serious concern, we cannot infer from these particular studies that participants had an inadequate grasp of random allocation and equipoise at the time they consented to participate.

Failure to understand about random allocation and equipoise could occur if trial information is just too complex for the patient to comprehend, or if the patient is not given sufficient time or opportunity to take it in. More accessible written information, or follow-up telephone conversations with a research nurse, might help with problems of this kind (e.g. Davis, Holcombe, Berkel, Pramanik & Divers, 1998; Edwards et al., 1998).

Written trial information will normally be supplemented with oral information, but it is the leaflet that is scrutinised by research ethics committees to check that it offers the necessary information for informed consent or refusal to participate in the trial. In the UK, guidelines have recently been established for approval of trial information leaflets by Multicentre Research Ethics Committees. Leaflets are expected to make a statement about equipoise, and a statement that treatments will be allocated at random. The particular wording suggested is "Sometimes because we do not know which way of treating patients is best, we need to make comparisons. People will be put into groups and then compared. The groups are selected by a computer that has no information about the individual, i.e. by chance. Patients in each group then have a different treatment and these are compared." There follows information about the chance of receiving each of the treatments (Committees, Ethics Committees (COREC), 2001).

Appelbaum and colleagues (1987; 1999); Appelbaum & Grisso, 2001) argue, however, that simply providing clear factual information is not sufficient to ensure patient understanding. They argue that patients hold a very strong assumption that their doctor acts in their best interests and offers what he or she considers is best for them, and patients are inclined to maintain or fall back on that assumption despite being given clear information that this will not happen in a RCT. According to this view, the solution is to ensure that patients realise they are entering a research study which does not follow the procedures of standard clinician-patient consultations.

Appelbaum et al's (1987; 1999) assumption that patients tend to interpret novel trial experiences in terms of their familiar framework of standard clinician-patient interactions is entirely in line with the long-standing view in cognitive and social psychology, according to which people are not passive recipients of information but rather are active interpreters who try to make sense of new input by drawing on their background knowledge and beliefs (e.g. DiSessa, 1996; Shank & Abelson, 1977). But Appelbaum and colleagues' approach apparently rests on the assumption that so long as patients abandon their inappropriate assumptions about treatment allocation in a trial setting, they are in a strong position to grasp and hold on to the information they are given about random allocation. Disabusing patients of incorrect assumptions may not however be sufficient to ensure they understand and accept the true situation. Evidence from the qualitative studies mentioned above (Featherstone & Donovan, 1998, 2002; Snowdon et al., 1997), suggests that at least some patients actively try to make sense of *why* randomisation is used. Their attempts may in some cases produce incorrect justifications and this may lead them to lose sight of the fact of initial equipoise. Some patients may fail to come up with any plausible reason for randomising, and may then lose sight of the fact of randomisation.

From the trialists' perspective, there are two intertwined justifications for randomisation, one ethical and the other scientific. The state of initial equipoise which motivates setting up the trial also provides the ethical justification for randomisation: it is acceptable to allocate participants to treatment arms at random only if relevant experts as a group have no convincing grounds for believing that one treatment arm has overall benefits compared with another. Note that the initial state of collective equipoise does not mean that potential trial recruits need themselves be in equipoise. A patient might legitimately have a preference for one treatment arm over another, based on his or her particular values in combination with accurate understanding of the current state of knowledge about the treatments under comparison (Ashcroft, 1999; Lilford & Jackson, 1995). Nor does the state of collective equipoise mean that individual clinicians need themselves be in personal equipoise; we come back to this point below.

The state of collective equipoise does not provide a sufficient justification for offering random allocation to treatment arms. The further necessary justification is the scientific one, that by minimising selection bias, a randomised trial is considered most likely to yield results that genuinely increase our knowledge (Schulz, Chalmers, Hayes, & Altman, 1995).

On the assumption that the UK recommended wording for approval by Multicentre Research Ethics Committees (given above) provides an example of current best practice, explanations of the scientific or ethical justifications for randomising in trial information leaflets may be either missing or obscure. The guidelines' recommended statement about equipoise is not explicitly presented as a justification for randomising, though it could be interpreted as a justification in the absence of any other, and it could be that at least some people find it an adequate justification. We examine this possibility in Studies 1 and 4 reported below.

Trial information leaflets, which provide only bald statements about randomisation and equipoise, may be adequate to permit a patient who has the relevant background knowledge to construct a true justification for randomisation. They may not be adequate, though, for many members of the lay public. The aim of the studies reported below was to investigate the background knowledge and assumptions that members of the public could bring to bear should they be invited to participate in a trial. We use hypothetical scenarios involving individual treatment and clinical and non-clinical research to examine the extent to which lay people consider randomisation to be justified given a state of equipoise (Studies 1 and 4), and the extent to which they consider random allocation to have scientific benefits over other methods of allocation (Studies 2–4). Since the ethical justification for randomisation relies on acceptance of the possibility of equipoise, we examined

lay views about the possibility of equipoise in Studies 1 and 4.

Background knowledge may be accessed differently in different contexts: medical vs. non-medical, and medical treatment vs. research. We included additional comparisons to examine such context effects. The rationale for these is given in the introduction to each study.

Methods common to Studies 1–4

Participants

Our 344 participants were adults attending a wide variety of part-time further education and leisure courses in 38 different classes at colleges in Staffordshire and South Cheshire, UK. Table 1 gives their ages, qualifications and occupations. No participant took part in more than one of the studies. Minimum sample sizes were determined a priori to allow us to detect medium sized effects with a power of at least 0.8 in the analyses using analysis of variance.

Procedure

With the permission of class teachers, one of the researchers (CK) gave a brief talk about the research to the participants in their classes, which varied in size from around 2 to 22 people. The researcher then handed out a single sheet to each participant with brief written scenarios and accompanying questions. When there were different between-subject conditions within one study (so participants were to receive one of several different sheets), we arranged the sheets in a fixed order and used random number tables to determine which condition should be first in the pile. This ensured that different sheets were distributed in roughly equal numbers to all classes who took part and that nobody had the same sheet as their immediate neighbour. Participants were asked to read the scenario and complete the questions individually. The sheets were handed out to everybody in the class with the instruction that anybody who preferred not to participate should hand in their sheet uncompleted at the end. Overall 95% of people invited to take part did so.

After the sheets had been filled in, the researcher asked for volunteers to talk through their answers individually to check on whether the scenarios and questions were interpreted as we intended and to gain an impression of the thinking that surrounded participants' written answers. The timing of classes and breaks determined how many volunteers could be interviewed from any one class, but we distributed the interviews across classes as much as possible. Interviewees were given a 5 pound voucher in recognition of the time spent. When all the data had been gathered, the

Table 1
Characteristics of participants in studies 1–4

		Study 1 (*n* = 82, 9 interviews)	Study 2 (*n* = 67, 8 interviews)	Study 3 (*n* = 67, 8 interviews)	Study 4 (*n* = 128, 19 interviews)
Age (yrs)	Range	21–61	21–81	20–80	19–82
	Mean	30.73	39.42	44.15	47.34
	SD	9.10	16.76	16.49	14.88
Gender	Male	20 (24.4%)	15 (22.4%)	11 (16.4%)	30 (23.5%)
	Female	62 (75.6%)	51 (76.1%)	48 (71.6%)	95 (74.2%)
	Missing	0 (0%)	1 (1.5%)	8 (11.9%)	3 (2.3%)
Occupation	Profess./manag./skilled non-manual	15 (18.3%)	8 (11.9%)	10 (14.9%)	44 (34.4%)
	Manual/semi-skilled/unskilled	16 (19.5%)	11 (16.4%)	8 (12.0%)	30 23.4%)
	Retired/student/no occupation	49 (59.8%)	47 (70.2%)	39 (58.2%)	48 (37.5%)
	Missing	2 (2.4%)	1 (1.5%)	10 (14.9%)	6 (4.7%)
Highest qualification	Degree	4 (4.9%)	7 (10.4%)	10 (14.9%)	39 (30.5%)
	Adv GCE (18 yrs)	19 (23.2%)	15(22.4%)	7 (10.4%)	14 (10.9%)
	GCSE (16 yrs)	42 (51.2%)	34 (50.7%)	19 (28.4%)	37 (28.9%)
	Other/none	11 (13.4%)	7 (10.5%)	19 (28.4%)	26 (20.3%)
	Missing	6 (7.3%)	4 (6.0%)	12 (17.9%)	12 (9.4%)

researcher returned to hold a debriefing session with each class. The studies were conducted in the sequence 1,2,4,3.

Study 1

We assessed participants' acceptance of the possibility of an individual doctor being in a state of complete uncertainty between two possible treatments, and their views about the acceptability of such a doctor suggesting deciding by chance which of two treatments to offer. It may be rare for individual trialists to be in equipoise when there is a state of collective equipoise that provides ethical justification for the trial (Alderson, 1996; Twomey, 1994). So far as we have been able to detect, though, patient information sheets do not make a distinction between individual and collective equipoise. Statements such as 'doctors do not know which treatment is best' can be interpreted as implying individual as well as collective equipoise. We decided therefore to find out whether participants accepted a move from individual equipoise to an offer of random

allocation, instead of attempting to explain the more complicated situation of collective equipoise without individual equipoise. Note that in this first study we asked about the *acceptability* of offering random allocation given uncertainty: there was no scientific rationale since there was no research context.

We invited participants to take the perspective of a patient who had consulted their doctor about a common and non life-threatening condition (severe back pain). We deliberately made the difference between the two possible treatments relatively trivial to minimise the chances that participants would form a preference for one treatment.

We were also interested in the extent to which participants' views were particular to a medical context. We created a parallel scenario involving a lawyer who could advise a client either to go to court or to settle out of court and in particular circumstances might be in equipoise about which was the better course of action. In both the medical and the legal scenarios a lay person consulted an expert professional about a matter of personal importance, and could reasonably expect that professional to act in their client's best interests. When

health and illness are involved though, people may be particularly inclined to seek or assume certainty rather than uncertainty. If so, the possibility of uncertainty, and of allocating on the basis of chance, might be better tolerated in the legal than in the medical context.

Method

Design. Participants received either a medical or a legal scenario.

Participants. See Table 1.

Procedure. Participants read one of two scenarios that provided either a medical or a legal context for professional uncertainty. The medical scenario asked participants to imagine they had consulted the doctor for severe back pain, and the doctor had explained there were two usual treatments, A and B. These were described briefly: one involved lots of small painless pulses within a treatment session and the other involved fewer but larger painless pulses. The two treatments had not been compared so the doctor did not know whether one was better, and only one could be used on any one patient. The legal scenario described a similar situation of uncertainty: participants were asked to imagine they had been injured in an accident and had consulted a lawyer about making a claim for compensation. The lawyer explained they could either settle out of court or go to court. The lawyer was sure that it was best to go to court if the offer was low, and better to settle if it was high. But the lawyer was unsure which was best if the offer was of a medium amount. After participants had read either the medical or the legal scenario, they read and gave yes or no answers to two questions:

'Do you think the doctor/lawyer could ever be completely unsure about which of two treatments is best/what's best, and truly not prefer one treatment/course of action over the other?'

'If the doctor/lawyer really was completely unsure, and did not prefer one over the other, would it be acceptable for the doctor/lawyer to suggest deciding at random? This would mean, for example, selecting a treatment /course of action by using a computer which has no information about the individual, i.e. by chance.' This final sentence was taken from the guidelines for trial information leaflets produced for UK Multicentre Research Ethics Committees (COREC, 2001), conformity or near conformity to which is now required by many UK medical ethics committees. We decided against elaborating on the meaning of random allocation, for example by adding an analogy such as tossing a coin. The results of a previous study indicated such clarification was unnecessary (Kerr et al., in press).

Table 2
Study 1: Frequencies of yes and no answers

Question	Medical scenario		Legal scenario	
	Yes	No	Yes	No
Can doctor/lawyer be unsure?	18	23	23	18
Suggest deciding by chance?	5	36	6	35

Results

The results are summarised in Table 2, which suggests the pattern of results is similar for the medical and legal scenarios. In both cases, many participants thought the doctor or lawyer could not be unsure, and a majority thought it was unacceptable to suggest deciding by chance. We checked these impressions with statistical analyses.

Medical vs. legal scenarios. For none of the comparisons that follow were there differences between the two scenarios, 95% confidence intervals (CIs) of the unpaired difference between scenarios spanned zero in all cases.

Acceptance of the possibility of uncertainty. For both the medical and the legal scenarios, participants were split as to whether or not the doctor/lawyer could be completely unsure. In both cases CIs spanned 0.5: medical scenario proportion "Yes" = 0.44, CI = 0.30 → 0.59; legal scenario proportion "Yes" = 0.56, CI = 0.41 → 0.70.

Suggesting deciding by chance. For both scenarios, significantly fewer than half the participants judged that it would be acceptable for the doctor/lawyer to suggest deciding by chance when he or she was completely unsure: medical scenario proportion "Yes" = 0.12, CI = 0.05 → 0.26; legal scenario proportion "Yes" = 0.15, CI = 0.07 → 0.28. A problem with interpreting this result is participants' unwillingness to accept that the doctor or lawyer could be completely uncertain; some "No" responders might have been denying the possibility of uncertainty rather than the acceptability of suggesting deciding by chance. However, amongst the subset of participants who accepted the possibility of uncertainty, only 2 out of 18 judged that it would be acceptable for the doctor to suggest deciding at chance, and 5 out of 23 judged that it would be acceptable for the lawyer to do so.

Qualitative data from interviews. The interviews suggested that participants understood the scenario and questions as intended. Comments about the legal and medical scenarios were remarkably similar. Table 3 contains quotes from interview participants illustrating the following themes. Some participants acknowledged

Table 3
Study 1: Illustrative quotes from interviews ($n = 9$)

Uncertainty (equipoise)	*Acknowledging uncertainty*: "He'll get results but he won't necessarily get all the feedback to him, will he, of the broader picture" (doctor) "Lawyers will never know everything, they will never know the entire story" (lawyer) *Uncertainty as ignorance*: "If they don't know they should send us to a specialist…who should provide that information" (doctor) "I'd have thought if he was totally unsure why not refer to somebody who was sure about what would be best for you?" (doctor) *Professionals should be certain*: "They must know quite a bit about each one…they should know which one would be best in your circumstances" (doctor) "I wouldn't think he would really be unsure about what's best, that's why you would go to them wouldn't you" (lawyer)
Deciding by chance (randomisation)	*Acceptable to decide by chance*: "I think it'll be all right to select at random seeing as they are both roughly the same sort of treatment" (doctor) *Unacceptable to decide by chance*: "I don't think chance decisions should be made with your health" (doctor) "I think it would be pretty unprofessional, I mean he's got other lawyers in his company that can advise… I mean taking a chance is probably too much" (lawyer)

individuals' uncertainty while others assumed that even if an individual did not know which course of action to take, somebody else was likely to know. Others assumed that a competent professional would or should know. Only one of our 9 interviewees thought it acceptable to suggest deciding by chance, the others found it unacceptable.

Discussion

None of our results suggested that participants thought about uncertainty any differently in the medical and legal contexts, and in both cases around half denied the possibility of individual equipoise. Given the suggestion above that trialists may often have a preference for one of the treatments, it could be argued that our participants' views were accurate. However, the interviews provided no sign that participants accepted collective equipoise and assumed individuals would have different views within that. Rather, the comments suggested that at least some participants assumed the required knowledge would or should be available, and the individual professional need only seek advice from others to reduce his or her uncertainty. Although the medical scenario stated that the two treatments had not been compared, participants apparently often failed to

realise or accept the implications of the absence of comparison.

Even more striking than the lack of acceptance of individuals' uncertainty was the view that it was not acceptable for the uncertain doctor to suggest deciding by chance. Our results indicate that potential trial participants are unlikely to accept that mere uncertainty, without any scientific context, provides sufficient grounds for randomising. In Study 4 we checked on this again in a slightly different way, by stating that the imaginary patient was equally willing to receive either treatment.

Study 2

In Study 2, we introduced a research context in which two treatments were to be compared. We were interested in whether, in the absence of any explanation or justification for randomising, participants recognized that random allocation has scientific benefits over allowing patients and doctors to choose which treatment to have. As in Study 1, we checked for context effects. This time our comparison scenario involved two ways of applying chemical treatment to sheep. Farmers could be allocated one of the two methods at random, or could be allowed to choose their method. The scientific benefits of

random allocation might be less easily recognised in the medical scenario, due to interference from the knowledge that doctors and patients normally choose the treatment that they think is most suitable.

Method

Design. Participants read both a clinical trial scenario and a sheep dip scenario, with order of presentation counterbalanced between subjects.

Participants. See Table 1.

Procedure. For the clinical trial scenario, participants read: 'Doctors sometimes have to make careful comparisons between two treatments in order to increase our knowledge about which is the best one. Below is an imaginary account of a typical comparison. We would like your views about how this should be carried out.' There followed the description of the two treatments, A and B, for back pain used in Study 1. The proposed study was then described: 'Doctors want to compare the two treatments in a scientific study. 500 patients with pack pain have agreed to take part. Half will be given treatment A and half will be given treatment B. This will be decided randomly, for example the treatments would be selected by using a computer which has no information about the individual, i.e. by chance.' Participants were then asked 'Once the study is done, how sure do you think doctors would be about which is the better treatment? Remember that treatments were allocated at random and neither doctor nor patient could choose which treatment a patient had.' Participants judged whether they thought doctors would be very sure (4), fairly sure (3), fairly unsure (2) or very unsure (1), or indicated that they did not know.

Next participants were asked 'If instead, the doctor and patient had chosen which treatment each patient was given, how sure do you think doctors would be about which is the better treatment?' Participants judged on the same scale as before.

The other scenario described scientists' comparison between two ways of preventing sheep from developing infections. One involved dipping the sheep in a trough of chemicals, and the other involved spraying the sheep with the chemicals as they passed through a plastic tunnel. Random allocation was described as in the trial scenario and participants judged how sure scientists would be at the end of the study about which treatment was the better, using the same scales as for the trial scenario. Finally, participants judged how sure the scientists would be if farmers had chosen which treatment to use on their farms.

Results

The results suggest that participants did not think that randomisation offered any advantage in the clinical context, though they may have done so in the context of sheep: the mean certainty score (sd) for the clinical trial scenario with random allocation was 2.74 (0.70), and with patient choice was 2.74 (0.83). For the sheep dip scenario the corresponding figures were 3.00 (0.63) and 2.62 (0.86). Statistical analysis confirms this impression. The certainty scores (1–4) were analysed using ANOVA with scenario (clinical trial or sheep dip) and allocation method (random or choice) as repeated measures, and order of scenarios (trial or sheep dip first) as a between group factor. Participants who indicated they did not know were omitted ($n = 4$). There were no significant main effects but the interaction between scenario and method was significant: $F(1,59) = 9.86$, $p = 0.003$. This was due to participants judging that with sheep treatments, scientists would be more sure with randomisation than farmers' choice, $t(60) = 2.646$, unadjusted $p = 0.01$, but with medical treatments there was no difference between judgements of how sure doctors would be with randomisation or choice.

Qualitative data from interviews. Interviewees' comments gave some indication of what they thought about the consequences of random allocation in comparison to letting patients and doctors choose their treatment. Table 4 contains illustrative quotes. A recurring theme was the idea that more information is taken into account with patient choice and so the outcome is likely to be better than with random allocation. One participant took the view that randomisation was no better than patient choice at controlling important variables. Some participants did however seem to judge that allocation or response bias could be avoided with random allocation.

Some comments about the sheep treatment scenario were very similar to those given in connection with the medical scenario: with choice, farmers could draw on their knowledge in order to achieve a better result, although as with the medical scenario, finding out which treatment was better was confused with achieving a better result for the individual sheep. With random allocation while again there were concerns of lack of control, some interviewees seemed to think this was less of a problem with sheep than with people. Comments about the sheep treatment scenario helped us interpret the finding reported above, that randomisation was judged to lead to greater certainty than allowing farmers to choose. Unexpectedly, some interviewees assumed that randomisation was the better method because if farmers could choose, they would all choose the sheep dip rather than the tunnel since they would then not have to invest in new equipment (though there was no information in the scenario to suggest this). There was no indication that participants appreciated the scientific reasons for randomising in the context of the sheep dip.

Implicit in two interviewees' discussions about controlling important variables was the idea of taking into

Table 4
Studies 2 and 3: Illustrative quotes from interviews ($n = 8$, $n = 8$)

Randomisation or deciding by chance	*Randomisation no better than patient choice*: [with randomisation or choice] "We don't know if they are all suffering from the same sort of back pain, there's too many uncertainties…I just think that the findings would be unreliable" (Study 2, medical) [with randomisation or choice] "They are not all being kept in identical conditions so there's too many other variants that could affect the outcome of the test… I think you need some control in there for the results to have any real bearing" (Study 2, sheep treatment) *Randomisation can reduce bias*: "You would get a good mixture" [but with choice] "That would be unsure because you would get so many patients want to do it one way and so many patients do it the other way, you wouldn't get a good mixture of the treatments" (Study 2, medical) "You're going to be very honest as to whether it's worked or not" [but with choice] "Some people have expectations, great the doctor says I'm going to feel better so I feel better now" (Study 2, medical) *Randomisation results in equal groups*: "they are going to like have been tested themselves each in equal amounts" [but with choice] "maybe one washing powder would have been chosen more than the other" (Study 3, washing powder)
Choice	*More information taken into account with choice than randomisation*: "Because the Dr would know the patients' situation, so he would be able to distinguish which one of the two treatments would probably work better in that situation so he would be able to clarify a bit more" (Study 3, medical) "They might have an understanding of what's best for them…" (Study 2, sheep treatment) "they know the situations don't they so…they'd know which one would probably be better for them" (Study 3, washing powder) *In light of no benefit from randomisation choice is better*: "at least you have the choice. It's better to have a choice than to have no choice at all" (Study 3, washing powder)
Differences between contexts	*Lack of variation or influence of sheep* (*Study 2*) "Sheep are sheep at the end of the day…I think because it is totally fifty-fifty on the treatment they would be very sure what the outcome would be" *Strong farmer preferences* (*Study 2*) "I'm looking at sort of financial reasons…to do things they have done in the past…. They might want to stick with the dip".

account key characteristics such as age, sex, severity and type of illness in assigning patients to treatment. The acceptability of such a form of assignment is explored further in Study 4.

Discussion

The results give no indication that participants appreciated the scientific benefits of random allocation of treatments in a clinical trial over patient choice. The most common response was to judge that doctors would be fairly sure which treatment was best with either method (50% of participants judged in this way for patient choice, and 46% did so for random allocation). Although the statistical analysis suggested that the benefits of random allocation were more likely to be appreciated for the sheep treatment scenario, the data from the interviews did not reveal understanding that random allocation avoids bias. On the contrary, they alerted us to the possibility that the result was due to participants making unforeseen assumptions about farmers' choices, and we followed up this possibility in Study 3.

Study 3

In this study we checked again on whether participants recognised the scientific benefits of random allocation, and on whether or not there were context effects.

Method

Participants. See Table 1.
Design and Procedure. The design and procedure were the same as for Study 2 except that we replaced the sheep treatment scenario with one involving a comparison of two washing powders.

Results

The results suggest that participants saw no advantage of randomisation in either context: for the clinical trial scenario the mean certainty score (sd) with random allocation was 2.95 (0.71), and with patient choice it was 3.10 (0.85). For the washing powder scenario the corresponding figures were 2.97 (0.72) and 2.90 (0.83). Statistical analysis confirms this impression. The certainty scores (1–4) were analysed using ANOVA with scenario (clinical trial or washing powder) and allocation method (random or choice) as repeated measures, and order of scenarios (trial or washing powder first) as a between group factor. Participants who indicated they did not know were omitted ($n = 5$). There were no significant main effects. The interaction between scenario and method approached significance: $F(1,56) = 3.54$, $p = 0.065$. If anything this reflected a different pattern from the significant interaction found in Study 2: there was a non-significant tendency for participants to give higher certainty scores with patient choice in the clinical scenario than in any of the other conditions.

Qualitative interview data. As in Study 2, the view that more information is taken into account with patient choice was voiced, though more interviewees in this group felt that the methods were equivalent (either just as sure or just as unsure with each). One participant concluded that choice was more desirable. Two participants saw an advantage of randomising in that it ensured equal-sized groups whereas choice might not, but they saw no other benefit. Two other participants raised the concern of response bias if patients/householders chose and believed their choice to be better, but no interviewees in this group considered randomisation to be advantageous in reducing allocation bias. Most participants judged the non-clinical scenario no differently from the clinical scenario. However, the washing powder scenario gave rise to three discussions similar to the two noted in Study 2 about key variables that might influence the outcome, and should therefore be taken into account, in this case age of washing machine, number, age and occupation of members of the household.

Discussion

For the clinical trial scenario, the results confirm those of Study 2: there was no sign that participants as a group thought a more certain result would be achieved by allocating treatments at random rather than by patient choice. We had wondered if participants would find it particularly hard to adopt a scientific perspective in a clinical context, and might acknowledge the scientific benefits of random allocation in a non-clinical context. This proved not to be the case with the washing powder scenario, in line with our suspicion that the quantitative result from the sheep treatment scenario was not due to understanding of the scientific benefits of randomisation.

Study 4

In our final study we clarified and built on our previous results. In Study 1 around half our participants were loath to accept that an individual doctor could be in equipoise. Comments made in interviews suggested that they sometimes assumed that a more expert colleague would know which treatment was best (even though it was stated that the treatments had not been compared), so in Study 4 we excluded that possibility by making explicit that there was no agreement amongst experts.

We also included a new condition in which the question about which treatment was best was the subject of a research study. Participants might more readily accept the possibility of equipoise in a research context that makes it legitimate, as compared to a treatment context in which uncertainty is perhaps an anomaly.

As in Study 1 we asked participants whether or not it would be acceptable for a doctor to suggest deciding at random which treatment to offer, but this time we stated that the patient was willing to receive either treatment. This wording avoided participants having to assume the doctor was completely unsure. Participants might judge random allocation to be more acceptable in a research context, and to find out we compared a research with a treatment context.

We built on the results of Studies 2 and 3 by examining again participants' recognition of the scientific benefits of random allocation. This time we compared their judgements of how sure experts would be with random allocation, with judgements of how sure they would be if matched groups were created. Comments made in the interviews of Study 2 and 3 prompted the creation of this new condition. Some interviewees saw the need to take into account patient characteristics

that might influence outcome. We wondered whether participants would recognise the scientific benefits of a design in which groups were deliberately matched on certain characteristics. If so, would allocation to such groups be judged acceptable even though, just as with random allocation, doctor and patient would have no control over the group to which a particular patient was allocated?

Method

Design. There were 8 groups: participants received either a back pain or an arthritis scenario, with either a treatment or a research context, and with questions about the knowledge gained with random allocation either preceding or following questions about the knowledge gained with matching.

Participants. See Table 1.

Procedure. Half the participants read a scenario focused on two possible treatments for back pain as used in Studies 1 and 2, and the other half read a scenario focused on two possible treatments for arthritis. One of the arthritis treatments involved taking two tablets in the morning, and the other involved taking one tablet in the morning and another in the evening. For both the back pain and the arthritis treatments participants were told, "The two treatments have not yet been compared and there is no agreement amongst the experts as to which one is better." This was intended to indicate a state of collective equipoise.

For the participants in the *treatment context* conditions, the subsequent procedure was as follows. Participants answered yes or no to the uncertainty question "Do you think the doctor could ever be completely unsure about which of the two treatments is best, and truly not prefer one treatment over the other?" This was followed by the acceptability question "Suppose you and your doctor agree that you would be equally willing

to receive either of the treatments, would it be acceptable for the doctor to suggest deciding at random?" (followed by clarification about allocating by computer as in Study 1).

For participants in the *research context* conditions these two questions were preceded by the statement "Your doctor asks if you would be willing to take part in a study comparing the two treatments. 500 people with arthritis/back pain will participate." The uncertainty question followed: "Do you think the doctor could ever be completely unsure ….?" The acceptability question was put within the context of the research study: "There are two ways of carrying out a study like this. In the first way you will be allocated to have one treatment, either A or B, at random…."

Following the acceptability question, participants given the research context then answered an additional question about the knowledge gained from the study: "If they allocate 250 patients to each of the treatments at random, once the study is done how sure do you think the experts would be about which is the better treatment?" As in Studies 2 and 3, participants gave a rating from 'very sure' (4) to 'very unsure' (1), or indicated that they did not know.

Participants given the research context also made acceptability and knowledge judgements about a second way of allocating treatments, by matching: "In the second way your details will be given to the research team who will ensure that the group of people receiving treatment A and the group of people receiving treatment B are well matched. They will allocate one treatment, either A or B, to you. Neither you nor your doctor will have any control over which treatment you receive. Would it be acceptable for the doctor to suggest deciding by matching?" Participants also judged how sure experts would be at the end of the study, as for random allocation. Half the research context participants made their judgements about acceptability of and

Table 5
Study 4: Frequencies of yes and no answers to questions about uncertainty and acceptability of allocation methods forming basis of comparisons

Question	Treatment context (N = 41)		Research context			
			Asked about randomisation first (N = 42)		Total including those asked about matching first (N = 87)	
	Yes	No	Yes	No	Yes	No
Can doctor be unsure?	18	23	23	19	46	40
Suggest deciding by chance?	10	30	18	23	42	41
Suggest deciding by matching?					64	21

Note: Cells that do not add up to the sample size have missing data.

knowledge gained from random allocation before their judgements about the matched design, and half had the reverse order.

Results

Table 5 shows the frequencies of yes and no answers to the questions about the doctor's uncertainty and acceptability of random allocation or matching. The table suggests that many participants in both the research and the treatment contexts thought the doctor could not be unsure, and many found it unacceptable for the doctor to suggest deciding by chance. There is a hint that randomisation may be less unacceptable in the research context. We checked these impressions with statistical analyses.

Acceptance of possibility of uncertainty in research and treatment contexts. To compare views about participants' acceptance of the possibility of uncertainty in treatment and research contexts we included only research context participants who had the random method first, since random allocation was the only method in the treatment context. In both the treatment and the research scenarios, participants were split as to whether or not the doctor could be completely unsure, in both cases CIs spanned 0.5: treatment scenario proportion "Yes" = 0.44 (CI = 0.30 → 0.59); research scenario (randomisation first) proportion "Yes" = 0.55 (CI = 0.40 → 0.68). Participants found it no easier to accept an individual doctor's uncertainty in a research than in a treatment context: CIs for unpaired differences spanned zero in arthritis and back pain scenarios considered separately, and for the two combined.

Acceptability of suggesting deciding by chance in research and treatment contexts. Again we compared the treatment context to the randomisation-first research context. In the treatment context significantly fewer than half the participants judged it acceptable for the doctor to suggest deciding by chance: proportion "Yes" = 0.25 (CI = 0.14 → 0.40). For the research context, participants were more split: proportion "Yes" = 0.44 (CI = 0.30 → 0.59). However, CIs for unpaired differences between treatment and research contexts spanned zero in arthritis and back pain scenarios considered separately, and for the two combined, providing no clear evidence that the two contexts differed in the acceptability of random allocation. Post hoc we included participants who had the matched design first, and comparing this larger sample of 83 for the research context with the treatment context, did find randomisation to be more acceptable in the research context (overall unpaired difference 0.26, CI = 0.07 → 0.41).

Acceptability of matching. Participants in the research context made judgements about the acceptability of a matched design as well as of randomisation. Thirty-four participants found both allocation methods acceptable,

14 found neither acceptable, 26 found only matching acceptable and 7 found only randomisation acceptable (missing data $N = 6$). Matching was judged more acceptable than randomisation: paired difference 0.23 (CI = 0.10 → 0.36).

Qualitative data about matching. Comments made in the interviews indicate why participants judged matching to be more acceptable than randomising. Thirteen interviewees received the research context and so had the opportunity to talk through their views about the matched groups design. Only 4 of these appeared to have interpreted the design as we intended, that is that two matched groups would be created and neither doctor nor patient would have control over the group to which an individual patient was allocated. In contrast, 7 of our interviewees appeared to have interpreted the design as matching treatments to patients, and 3 of these judged matching to be more acceptable than random allocation for that reason: "I think you'd have a better success with this one, you and the doctor actually getting together and deciding together if it could work for you" (note also the focus on individual treatment success rather than gain in scientific knowledge); "Yes that's taking into account their medical history, their age, weight whatever, everything you know as a doctor about your patient and saying yes this probably would be the best one." For the remaining 2 of the 13, it was unclear how they had interpreted matching.

Knowledge gained with randomising and matching. Participants given the research context made judgements about the knowledge gained with random allocation and with matching. The mean certainty score (sd) with random allocation was 2.64 (.74), and with matching it was 2.77 (0.64). Certainty ratings (1–4) were analysed by ANOVA, with design (random allocation vs. matching) as a repeated measure and order (random or matching first) as a between subjects variable. Participants who judged that they did not know ($n = 15$) were excluded from this analysis. The ANOVA showed no significant main effects or interactions. The results give no indication that participants discriminated between random allocation and matching in terms of the knowledge derived from the study. For both methods, the majority of participants judged that experts would be fairly sure: 46 out of 85 (54.1%) with randomisation and 54 out of 86 (62.8%) with matching. Note, though, that the knowledge judgements need to be interpreted taking into account the information mentioned above, that some participants apparently treated the matching design as equivalent to patient choice.

Discussion

We confirmed our previous finding that around half the participants are loathe to accept that a doctor might genuinely not know what treatment is best. In this study

we made explicit that there was a state of collective equipoise, and participants again remained split as to whether or not an individual doctor could be in equipoise. It made no difference when the state of uncertainty was within a research context.

As in Study 1, more than half the participants thought it unacceptable for the doctor to suggest deciding by chance in a treatment context. There was a hint that randomisation might be less unacceptable in a research context, though even then only around half the participants judged it to be acceptable.

As in Studies 2 and 3, participants seemed to be insensitive to the advantages and disadvantages of different allocation methods for achieving an advance in knowledge. In Studies 2 and 3 they did not judge that more knowledge would be derived from a study involving random allocation than from one in which patients chose their treatment. In Study 4 they judged random allocation to be no more informative than matching. At least some participants, though, appear to have interpreted matching as being equivalent to patient choice, so judgements about knowledge gained need to be interpreted within that context.

The same proviso applies to the finding that participants judged allocation to matched groups to be significantly more acceptable than random allocation. Although only a small sub-sample of the participants talked through their answers for us, the high incidence of clear misunderstanding of the matched design argues against accepting the quantitative result at face value. We strongly suspect that the higher acceptability ratings difference can be accounted for by participants wrongly assuming that treatments would be matched to patients' needs.

This unexpected interpretation of the matched groups design is itself of interest. It is in line with Appelbaum et al's (1987) argument, summarised in the introduction, that trial participants are inclined to fall back on their assumption that treatments are selected in accordance with patients' needs. Even though our scenario stated "Neither you nor your doctor will have any control over which treatment you receive", some (but not all) participants seemed to overlook this. We come full circle: the very problem which prompted this research intruded even when we encouraged participants to take a scientific perspective on the trial, rather than the perspective of a patient receiving treatment within the trial.

Final discussion and conclusions

This research arose from the considerable body of evidence that despite efforts to simplify language and otherwise clarify trial information, participants in RCTs seem still to be at risk of failing to grasp, or of losing sight of, information that allocation to treatment arms is at random, and about the initial state of equipoise. Our results highlight three core issues that could be starting points for misunderstanding:

First, many of our participants were reluctant to accept that an individual clinician could be completely unsure about which of two treatments was better. This was apparent in Study 1 where participants made judgements in a treatment (rather than a research) context, and in Study 4 in which they made judgements in both treatment and research contexts. Potential trial recruits might find that proffered information about equipoise conflicts with their belief that clinicians will hold treatment preferences. Perhaps this belief is accurate; perhaps many trialists do hold weak (or even strong) treatment preferences within the context of collective equipoise. But so far as we can ascertain, statements of equipoise in trial information leaflets make no distinction between collective and individual equipoise. Participants who are given the information that "doctors do not know which treatment is best" are unlikely to work out for themselves that although individual doctors may have preferences there is no consensus view that one treatment is better than the other. Further research might investigate what degree of uncertainty amongst individual clinicians the lay public is ready to accept in various clinical circumstances, and whether people think an individual clinician's preference should be revealed and used as a potential basis for treatment.

Second, participants often found it unacceptable for a clinician to suggest deciding between treatments at random. This was evident in Study 1, and in Study 4 when the patient was equally willing to receive either treatment. Participants might have accepted that when current knowledge gives no grounds for choosing between treatments, and the patient has no preference, nothing is lost if chance decides which one the patient has. This view was certainly not prevalent.

Third, participants assumed that just as much knowledge would be gained about which treatment is better if patients and doctors chose their treatment, rather than if treatments were allocated at random. We found this in Studies 2 and 3. In Study 4 participants did not discriminate between random allocation and a matched design in terms of the knowledge gained, though the matched design was sometimes misinterpreted as involving patient choice. These results reveal a clear discrepancy between the assumption underlying trial design, that randomisation maximises the knowledge gained from the trial, and the assumption of many participants that randomisation holds no scientific advantage.

Taken together, these results render it unsurprising that many recruits to RCTs apparently fail to make sense of descriptive trial information about equipoise

and randomisation. Importantly, even if potential recruits *did* fully understand the initial state of collective equipoise and the scientific rationale for randomising, they might still have their own preferences for one treatment and for this or other valid reasons choose not to enter the trial. Furthermore, they might still think it wrong to offer random allocation: understanding the perspective of the trialists does not require accepting it as valid. Even amongst clinicians whose understanding is certainly not in doubt, different views are held about the relative value of knowledge gained from randomised and non-randomised trials (e.g. Abel & Koch, 1999; Herman 1998; Pullman & Wang, 2001). Our results suggest that clear, descriptive trial information may not permit potential trial recruits to hold such an informed view.

A next step is to ascertain the consequences of providing potential recruits with an accessible explanation of the scientific benefits of random allocation given collective equipoise, and this is currently being investigated. We also need to consider how far our results can be generalised to potential recruits to real clinical trials. Insofar as we diagnosed accurately the background assumptions of our participants, we have a useful picture of the assumptions that many potential trial recruits could draw upon. If anything, our participants had a better chance of reflecting from a scientific perspective, given that they had no personal involvement. Any mismatch between our participants' assumptions and those underlying trial design is therefore unlikely to be avoided in a real trial unless additional explanatory information is provided. The consequences of providing such information, though, may be different in hypothetical studies and in real trial settings and it may be at that point that particular caution needs to be exercised in generalising from hypothetical studies.

Acknowledgements

References

Abel, U., & Koch, A. (1999). The role of randomization in clinical studies: Myths and beliefs. *Journal of Clinical Epidemiology*, *52*(6), 487–497.

Alderson, P. (1996). Equipoise as a means of managing uncertainty: Personal, communal and proxy. *Journal of Medical Ethics*, *22*, 135–139.

Appelbaum, P. S., & Grisso, T. (2001). *MacCAT-CR: MacArthur competence assessment tool for clinical research*. Sarasota, USA: Professional Resource Press.

Appelbaum, P. S., Grisso, T., Frank, E., O'Donnell, S., & Kupfer, D. J. (1999). Competence of depressed patients for consent to research. *American Journal of Psychiatry*, *156*, 1380–1384.

Appelbaum, P. S., Roth, L. H., Lidz, C. W., Benson, P., & Winslade, W. (1987). False hopes and best data: Consent to research and the therapeutic misconception. *Hastings center report*, *17*(2), 20–24.

Ashcroft, R. (1999). Equipoise, knowledge and ethics in clinical research and evidence. *Bioethics*, *13*(3/4), 314–326.

Central Office for Research Ethics Committees (COREC). (2001). *Guidelines for researchers: Patient information sheet and consent form*. Available at www.corec.org.uk

Cochrane Controlled Trials Register. (2002). Available at www.nelh.nhs.uk/cochrane.asp

Davis, T. C., Holcombe, R. F., Berkel, H. J., Pramanik, S., & Divers, S. G. (1998). Informed consent for clinical trials: A comparative study of standard versus simplified forms. *Journal of the National Cancer Institute*, *90*(9), 668–674.

DiSessa, A. A. (1996). What do just plain folk know about physics? In D. R. Olson, & N. Torrance (Eds.), *The handbook of education and human development* (pp. 12–12). Oxford, UK: Blackwell.

Edwards, S. J. L., Lilford, R. J., Braunholtz, D. A., Jackson, J. C., Hewison, J., Thornton, J. (1998). Ethical Issues in the design and conduct of randomised controlled trials. *Health Technology Assessment*, *2*(15).

Ellis, P. M., & Butow, P. N. (1998). Focus group interviews examining attitudes to randomised trials among breast cancer patients and the general community. *Australian and New Zealand Journal of Public Health*, *22*(5), 528–531.

Featherstone, K., & Donovan, J. L. (1998). Random allocation or allocation at random? Patients' perspectives of participation in a randomised controlled trial. *British Medical Journal*, *317*, 1177–1180.

Featherstone, K., & Donovan, J. L. (2002). Why don't they just tell me straight, why allocate it? The struggle to make sense of participating in a randomised controlled trial. *Social Science and Medicine*, *55*, 709–719.

Freedman, B. (1987). Equipoise and the ethics of clinical research. *New England Journal of Medicine*, *317*(3), 141–145.

Glogowska, M., Roulstone, S., Enderby, P., Peters, T., & Campbell, R. (2001). Who's afraid of the randomised controlled trial Parents'? views of an SLT research study. *International Journal of Language and Communication Disorders*, *36*(Suppl.), 499–504.

Herman, J. (1998). Shortcomings of the randomized controlled trial: A view from the boondocks. *Journal of Evaluation in Clinical Practice*, *4*, 283–286.

Hietanen, P., Aro, A. R., Holli, K., & Absetz, P. (2000). Information and communication in the context of a clinical trial. *European Journal of Cancer*, *36*, 2096–2104.

Johnson, N., Lilford, R. J., & Brazier, W. (1991). At what level of collective equipoise does a clinical trial become ethical? *Journal of Medical Ethics*, *17*, 30–34.

Kerr, C. E. P., Robinson, E. J., Stevens, A. J., Braunholtz, D. A., Edwards, S. J., Lilford, R. J. (in press) Randomisation in Trials: Do potential trial participants understand it and find it acceptable? *Journal of Medical Ethics.*

Lilford, R. J., & Jackson, J. (1995). Equipoise and the ethics of randomisation. *Journal of the Royal Society of Medicine*, *88*, 552–559.

Pullman, D., & Wang, X. (2001). Adaptive designs, informed consent, and the ethics of research. *Controlled Clinical Trials*, *22*, 203–210.

Schulz, K. F., Chalmers, I., Hayes, R. J., & Altman, D. G. (1995). Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *Journal of the American Medical Association, 273*(5), 408–412.

Shank, R. C., & Abelson, R. P. (1977). *Scripts, goals and understanding.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Snowdon, C., Garcia, J., & Elbourne, D. (1997). Making sense of randomisation; responses of parents of critically ill babies to random allocation of treatment in a clinical trial. *Social Science and Medicine*, *45*(9), 1337–1355.

Twomey, J. G. (1994). Investigating paediatric HIV research ethics in the field. *Western Journal of Nursing Research*, *16*(4), 413–414.

van Stuijvenberg, M., Suur, M. H., de Vos, S., Tjiang, G. C. H., Steyerberg, E. W., & Derksen-Lubsen, G., et al. (1998). Informed consent, parental awareness, and reasons for participating in a randomised controlled study. *Archives of Disease in Childhood*, *79*, 120–125.