An analysis of decision letters by research ethics committees: the ethics/scientific quality boundary examined

E L Angell,1 A Bryman,2 R E Ashcroft,3 M Dixon-Woods1

ABSTRACT

Objectives: The performance of NHS research ethics committees (RECs) is of growing interest. It has been proposed that they confine themselves to "ethical" issues only and not concern themselves with the quality of the science. This study aimed to identify current practices of RECs in relation to scientific issues in research ethics applications.

Methods: Letters written by UK RECs expressing provisional or unfavourable opinions in response to submitted research applications were sampled from the research ethics database held by the Central Office for Research Ethics Committees. Ethnographic content analysis (ECA) was used to develop a coding framework. QSR N6 software was used to facilitate coding.

Results: “Scientific issues” were raised in 104 (74%) of the 141 letters in our sample. The present data suggest that RECs frequently considered scientific issues and that judgments of these often informed their decisions about approval of applications. Current processes of peer review seemed insufficient to reassure RECs about the scientific quality of applications they were asked to review.

Conclusions: This study provides evidence that scientific issues are frequently raised in letters to researchers and are often considered a quality problem by RECs. In the discussion, the authors reflect on how far issues of ethics and the policy implications.

The practice of ethical review in health-related research has become increasingly contentious, with recent interventions in the debate arguing that ethical review of most health services research is fatuous.1 In the UK, the performance of research ethics committees (RECs) has been widely criticised. Excessive demands and inappropriate conservatism on the part of the RECs have been blamed for obstructing or impeding important research that would be of benefit to patients.2 These charges have joined more longstanding complaints about the apparent capriciousness and inconsistency of RECs.3-5 The perceived growth in the influence and scope of ethical review has been a source of much comment.6 A particular and growing focus of criticism now concerns the perceived inclination of ethical review to stray into the territory of science, rather than confining itself to more traditional ethical concerns such as not harming participants, informed consent, privacy and avoiding deception of participants.7

At present, under the UK Research Governance Framework,8 responsibility for ensuring the quality of science technically rests with the so-called “sponsor” of any study. Protocols submitted for ethical review should, under the research governance arrangements, have already been peer-reviewed and had prior critique by methodological experts. The Governance Arrangements for Research Ethics Committees (GAfREC) states that:

“It is not the task of an REC to undertake additional scientific review, nor is it constituted to do so, but it should satisfy itself that the review already undertaken is adequate for the nature of the proposal under consideration.”

However, GAfREC also commends that RECs be “adequately reassured” about particular aspects of the scientific design and conduct of the study, thus offering discretion to RECs about how far they accept prior assessments of the quality of the science. The question of whether RECs should concern themselves with scientific matters has been raised in several countries worldwide. In a North American context, the term “ethics creep” has been devised to denote:

“a dual process whereby the regulatory structure of the ethics bureaucracy is expanding outward, colonizing new groups, practices, and institutions, while at the same time intensifying the regulation of practices deemed to fall within its official ambit.”

In the UK, the issue has been given renewed emphasis by the recent Report of the ad hoc advisory group on the operation of NHS research ethics committees,9 which stated that it did “not believe that RECs should function as a secondary form of scientific review” and that “RECs should deal with ethical rather than scientific review”. These recommendations raise important questions about how far RECs should assume that the quality of research ethics applications is assured through the processes of peer review that are supposed to have occurred before their deliberations. In addition, some have argued that the science/ethics distinction is incoherent, and that RECs should be permitted to consider the science and are under an obligation to do so.10

However, there is little systematic evidence about how far RECs currently do concern themselves with scientific issues, which are the issues that are especially likely to engage their attention, and what might be the consequences if they were to suspend judgment on the science. In this paper we report an analysis of letters composed by NHS
The letter conveyed a “provisional” or “unfavourable” opinion (64% of decisions at first review, and require applicants to make a response to the REC addressing issues raised in the letter before a final opinion can be issued. The final opinion may be either favourable or unfavourable. An “unfavourable” opinion (~8% of all submissions) at first review amounts to a rejection. Researchers have the option to either resubmit a new application (taking into account the issues raised) or to appeal (in which case no changes can be made to the documentation).

Some applications are withdrawn (~10% before review by a REC (eg, because the researchers have decided not to proceed), 3% after a provisional opinion has been issued). RECs may also decide that applications are “outside remit” or that advice should be sought from an external expert (such as a methodologist or specialist clinician) before giving a formal opinion. *Data based on the period October 2005–March 2006.

Box 1: Decisions RECs may make

- A “favourable” opinion means that an application is approved without further amendments; these constitute ~15% of decisions made by RECs at first consideration of an application.
- “Provisional” opinions constitute ~64% of decisions at first review, and require applicants to make a response to the REC addressing issues raised in the letter before a final opinion can be issued. The final opinion may be either favourable or unfavourable.
- An “unfavourable” opinion (~8% of all submissions) at first review amounts to a rejection. Researchers have the option to either resubmit a new application (taking into account the issues raised) or to appeal (in which case no changes can be made to the documentation).

RECs in response to applications by researchers, aiming to identify which types of scientific issue, if any, RECs identify as troubling them in their ethical review of applications.

METHODS
Standard Operating Procedures issued by the Central Office for Research Ethics Committees (COREC) require that each REC in the UK register the applications it reviews on the national Research Ethics Database (RED). We were granted access to the RED by COREC for the purposes of our project. Our project included letters written by RECs to applicants following the first meeting at which an application was considered (ie, we did not include letters arising from consideration of researchers’ responses to amendments previously requested by the REC). Three possible decisions can be made by RECs when they first consider applications: favourable, provisional and unfavourable (Box 1). We were interested in applications that received an unfavourable or provisional decision at first review, as such decisions indicated that there was an issue in the application that troubled the REC. For applications in our study that received a provisional opinion, the final decision was recorded where available (a few decisions were unavailable on the database).

Criteria for inclusion of a letter in our sample were as follows:
- The letter conveyed a “provisional” or “unfavourable” opinion.
- The letter was about an application considered by a REC for the first time during our “eligible periods”: July 2005, October 2005, January 2006 and April 2006. These periods were chosen to minimise seasonal effects in application submission.

The 55 RECs that did not upload letters to the RED during our eligible periods were excluded, leaving 115 RECs from which letters could be sampled. The first letter that met our eligibility criteria for each of the 115 RECs was chosen for inclusion in the study. Unfavourable opinions were purposefully over-represented to yield sufficient letters for analysis, so that they formed 20% of the initial sample. The remaining 80% of letters were those applications with provisional opinions. In addition, because applications that initially received a provisional decision but were subsequently issued with an unfavourable opinion were of particular interest, this type of application was also purposefully over-sampled by including all such applications between March 2004 and July 2006 for which a letter was available. More than one “provisional” letter from some RECs was therefore included in the study.

Descriptive information about each application—for example, clinical drug trial, qualitative study, student project—was recorded. We also recorded whether the proposal had been peer-reviewed prior to submission to the REC. There are several different classes of peer review (ranging from review within the project team to external independent review), and tick-box categories for these have changed over time on the REC form. We report the categories ticked by applicants when they submitted their application.

To analyse the letters, we used ethnographic content analysis (ECA). ECA requires the development of a coding scheme or framework grounded in the data. The framework was generated initially through close inspection and comparison across the texts of letters used a previous study but was modified extensively in response to the new data in this project. Explicit specifications were devised to aid data assignment, which was facilitated by the use of QSR N6 software. Care has been taken to anonymise quotations from the letters and where appropriate identifying details have been modified or removed.

RESULTS
Table 1 shows the outcomes of the 141 letters that were selected for inclusion in our study; 23 applications were given an unfavourable opinion at first review and 118 were issued with a provisional opinion. Of these provisional opinions, 85 were later issued with a favourable opinion and 20 with an unfavourable opinion. Five were withdrawn by the researcher, one was withdrawn by the REC and one was still awaiting a response from the researcher at the time we sampled (deemed “final decision unknown”). A mix of application and applicant types was achieved in our sample. Of applications for which the final decision was known, 19/40 (48%) of applications to undertake intervention projects were approved, 56/56 (64%) of non-intervention projects were approved and 30/38 (79%) qualitative studies were given a favourable final opinion.

Scientific issues were raised in a large proportion (n = 104, 74%) of the 141 letters. Though any quantitative analysis of these data will necessarily be tentative, given our sampling strategy, there is some evidence that RECs’ judgment of scientific issues were decisive in influencing the outcome of applications. “Troubles” regarding scientific quality were raised in 51 (60%) of the 85 applications that were initially provisional.

| Table 1 | REC letters and outcomes of the ethical review process | | | |
|---|---|---|---|---|
| **Full decision** | **Total letters** (n) | **Intervention studies** (n (%)) | **Non-intervention studies** (n (%)) | **Qualitative studies** (n (%)) |
| Provisional then favourable | 85 | 19 (22) | 36 (42) | 30 (35) |
| Unfavourable at first review | 23 | 8 (35) | 11 (48) | 4 (17) |
| Provisional then unfavourable | 26 | 13 (50) | 9 (35) | 4 (16) |
| Provisional, final decision unknown | 7 | 3 (43) | 4 (57) | 0 (0) |
| **Total letters** | 141 | 43 (30) | 60 (43) | 38 (27) |

*Data based on the period October 2005–March 2006.*
and later favourable. However, issues regarding scientific quality were raised in 24 (92%) of the 26 applications that were unfavourable at first review. Scientific issues were also raised as troubles in all 23 applications that were unfavourable at first review, suggesting that issues of scientific quality were strongly associated with applications that were initially or eventually deemed unfavourable.

Within the overall category of “scientific issues”, we generated nine subcategories to characterise the types of issue raised by RECs (table 2): the sample; choice of methods; the research question; the measuring instrument; analysis; bias; feasibility; equipoise; and “other” design issues. Our analysis focuses on the 104 letters in which scientific issues were raised.

### Sampling

Issues relating to sampling were raised in 68 (65%) of the 104 letters in which scientific issues were raised by RECs. The most common trouble relating to the sample concerned inclusion and exclusion criteria (19 letters). RECs frequently requested more information, asked researchers to exclude certain groups of people, sought justification for the inclusion of particular vulnerable groups, or queried how potential participants would be identified. A concern for sampling criteria to be transparent and for a full justification of inclusion and exclusion criteria was prominent.

“The committee felt that the exclusion criteria for the study although sensible were not felt to robustly exclude subjects at risk of a severe reaction based upon experience to date. The committee could not identify alternative criteria that would fulfil this requirement.” (Letter 67, provisional opinion, review within institution)

Eleven letters expressed concern about the size of the sample. They requested clarity, justification of the sample size or for calculations to be re-done, or suggested that the sample size might be too small, especially when derived without the help of a statistical expert.

“The power calculation of 45% was too low and would not identify any real difference. The purpose of conducting the study is, therefore, under question.” (Letter 137, provisional opinion, review within company, internal review)

### Choice of methods

Issues relating to choice of methods were raised in 52 (50%) of the 104 letters in which scientific issues were raised. Most commonly, RECs expressed concern that the rationale for the methods was unclear (27 letters).

“There was significant confusion over the title and the design throughout the application […]” (Letter 24, unfavourable opinion, independent external review, review within institution)

RECs often made suggestions for alternative ways of designing studies (17 letters), including the usefulness of control groups, the sources of tissue samples or data, or randomisation.

“Following discussions with you we thought that the control group would not be helpful given the large number of variables and that to treat one disease state using the research subjects as their own controls would help you achieve at least some of the answers you were trying to obtain.” (Letter 10, provisional opinion, independent external review, internal review)

### Research question

In 29 letters, there were issues relating to the research question. The most common concern was lack of clarity (15 letters).

“There appears to be some confusion about the status of this study. Although it is presented as a pilot study, statistical advice given for a similar study will be used in this present study (A45-2), which indicates that this is not actually a pilot study. Furthermore the research questions do not appear to be those of a pilot study.” (Letter 90, unfavourable opinion, review within institution)

Other concerns relating to the research question included queries about why the study was being undertaken (4 letters), suggestions for alternative research questions (4 letters), questions about whether the study would produce meaningful results (3 letters) and concern that the research question might be too ambitious or complex (2 letters).

“Members suggested that it was preferable to do the research using routinely collected blood samples and simplifying the research question.” (Letter 42, unfavourable opinion, review within institution)

### Measuring instruments

Queries or concerns about measuring instruments—for example, questionnaires and interview schedules—were raised in 28 (27%) of the 104 letters.

“Question 15 provides only negative responses. The committee suggested taking advice from the Clinical Psychologist in order to suggest some neutral/positive responses.” (Letter 22, provisional opinion, review within institution)
The usefulness of the measures to be used (10 letters), rationale (8 letters) and the ability of the study design to answer the research question (8 letters) were also questioned.

“The Committee considered that the study will not achieve the research question. The study design will only test the cream. This would leave the placebo patients denied a proven treatment for an investigation with no clear purpose.” (Letter 33, unfavourable opinion, no peer review)

**Data analysis**

RECs were concerned about issues relating to the analysis of data in 23 (22%) of the 104 letters, and in 10 letters they expressed specific concerns about statistical analysis.

“The Committee noted that the application provided no information whatever on the statistical analysis that would be undertaken, and I should be grateful if you could provide clarification and confirmation from a Statistician independent of the study of the validity of the proposed calculations.” (Letter 74, provisional opinion, review within institution)

**Bias**

RECs queried aspects of the study design that might bias the findings in 16 (15%) of the 104 letters. The most common concern related to the relationship between the researcher and participants (8 letters), but other issues relating to the potential for bias resulting from the design of the study were also raised.

“You indicated that financial limitations prevented you from undertaking transcription verbatim. The question came up whether being selective might lead to errors arising.” (Letter 94, provisional opinion, internal review)

**Feasibility**

The question of how likely it was that the work proposed by the researchers would be feasible was raised by RECs in 12 (12%) of the 104 letters. In six letters, RECs were concerned that recruitment might be slower or more difficult than anticipated. Other issues included burdens on staff, methods for extracting data, or the suitability of the research site and competence of the researchers.

“Members thought that the methodology used to recruit participants was unrealistic and as such recruitment could be a problem, as GPs may not find the time to distribute the information sheets, especially if the study was unlikely to produce significant results. Members strongly recommended that an alternate [sic] methodology should be used.” (Letter 42, unfavourable opinion, review within institution)

**Equipoise**

Concerns about lack of equipoise were raised in 10 (10%) of the 104 letters. Such concerns included assumptions about preferences (2 letters), assuming that one treatment is better than another (1 letter), not verifying a new method against an established method (1 letter), the effect of vulnerable groups (1 letter) and assuming that all discrimination is negative (1 letter).

**Other study design issues**

Other design issues were raised by RECs in 27 (26%) of the 104 letters. These mainly related to scientific peer review (16 letters)—that it had not been submitted, that it was inadequate or that the researcher should address the concerns therein.

“An independent external scientific critique specific to this area of expertise is required, the review that was submitted was considered inadequate by the Committee.” (Letter 126, provisional opinion, independent external review, internal review)

Other unspecified issues relating to study design included concerns of a general nature (4 letters), missing or incorrect information (4 letters), issues relating to research governance approval (3 letters) and data monitoring (5 letters).

“The Committee expressed unhappiness about the scientific presentation of this study and feels that it needs significant revision.” (Letter 6, unfavourable opinion)

**DISCUSSION**

Our analysis suggests that REC letters frequently raised issues of “science” and sought clarification or amendment of methodological issues. Sampling appeared to be an area that was especially likely to be a focus of concern in REC letters, but many other categories of scientific “trouble” seemed to be important to RECs, and these may influence the decisions made by the RECs. These findings suggest that RECs do not seem to find sufficient reassurance about the quality of science from peer review conducted before applications are seen by RECs.

Our study does have several important limitations. In particular, we did not analyse the applications themselves, only the letters written in response. Our sample aimed to represent different types of decision and is not fully representative of all types of application. Nonetheless, our analysis does suggest that clarification of the case for regarding matters of research quality as ethical issues is needed, as is consideration of the limits and extent of the role of RECs in this area.

Clearly, one way of explaining our finding that RECs often concern themselves with matters of science is to treat it as evidence of “ethics creep” and territorial expansion. Other explanations should, however, be considered. One problem for RECs, for example, is how they can be assured that the scientific review carried out before they see an application is adequate. At present, it is not at all clear how RECs should satisfy themselves that the application has undergone appropriate review, since the current application form requires applicants to state what kind of scientific review has been undertaken but not necessarily to include the reports with their application.

The problem of being assured of the quality of prior review is of particular importance to RECs, because our data suggest that RECs tend to see research as a context in which the quality of research, considered broadly, has ethical implications. RECs, faced with what they consider to be a scientifically poor or dubious project, are confronted with the dilemma that such studies may pose risks. Poor quality health research, may, for instance, be harmful to future patients, whose treatment could be based on inadequate or misleading evidence, and unfair to present research participants, in that they are subject to the risks and inconveniences of participation in unworthy research. RECs may therefore feel that they are entitled to concern...
themselves with issues of science, on the grounds that bad science is bad ethics.

In this respect, RECs may have a particular concern with ensuring that patients taking part in studies are not harmed. In the detailed analysis of the issues the REC letters in our sample were raising were many issues that were concerned with the direct effects of poor study design on the well-being of people, such as the possibility that people in the placebo arm of a trial would be denied a known effective treatment (the cream mentioned in letter 53, for example) for no clear benefit. Such examples legitimate the interest of RECs in scientific issues and strengthen the argument that RECs should adjudicate on these matters.

Clear examples such as this should not, however, obscure the fact that there were many examples in our dataset in which there was unlikely to be scientific consensus about the issue at hand—for example, in relation to construction of questionnaires or sample size. The parlaying of such issues into matters of “harm” is arguably more problematic. As all those in the research community are aware, referees of research proposals often vary in their understanding of what is “good science”, and scientific review conducted within the context of a REC system is unlikely to be any different: there is no external infallible scientific authority to which appeals might be made. There is thus the potential that RECs might reject applications on scientific grounds that within the scientific community of practice in which they originated would be regarded as having satisfactory study designs.

A related process is that RECs may parlay scientific issues into harm-related form in order to deem them “ethical” issues, and this may be because RECs are concerned with issues of fairness. Fairness is problematic in ethical review for two reasons. First, issues of fairness are often debatable and obscure, whereas issues of harm are more clear-cut. Second, although researchers are under clear obligations not to harm their subjects, it is not clear that they are obliged to treat participants and non-participants fairly. Our analysis revealed a widespread concern among RECs about issues of sampling, and inclusion/exclusion criteria. Thus a scientific issue (exclusion/inclusion criteria) can be redefined as an ethical one (protecting individuals from unfair and harmful discrimination).

Any account of the distinction between science and ethics must also recognise the more general problem of distinguishing between science and non-science. To require RECs to deny themselves consideration of scientific issues, one has to accept an unambiguous distinction between ethics and science. The evidence we present here suggests that it is difficult to sustain such a distinction in practice, even if it is available in theory. Our data suggest that the science/ethics distinction can be seen as the outcome of a social process, rather than an a priori conceptual distinction. As Gieryn’s analysis of boundary work has shown, philosophers and sociologists of science have long struggled with the problem of “demarcation”: how to identify the unique and essential characteristics of science. Gieryn identifies boundary work as a rhetorical effort that involves the attribution, by scientists, of selected characteristics to the institution of science, for purposes of constructing a social boundary that distinguishes “science” from “non-science”. On this account, attempts to protect the science from the criticism of RECs involve boundary work rather than resting on an unchallengeable and uncontested distinction between science and ethics.

Our findings raise questions about the appropriate policy response. Better assurance of the quality of the science might reduce the potential for conflict between RECs and researchers. One response might be to focus on improving the quality of peer review before applications reach the committee stage and to ensure that all scientific review reports are made available to committees, the better to discourage RECs from considering scientific issues. One possibility might be that some types of peer review are considered more authoritative than others—for example, in the UK, research proposals that have been through certain funders will already have been subjected to rigorous peer review. It should be recognised, however, that prior scientific review by funders may not guarantee that all of the issues that our data suggest are of concern to RECs have been reviewed in detail; some funders require only brief details of such issues, and project specifications may change between approval by the funder and submission to the REC. Any system that relies on improved peer review before submission to the REC would also require caution to avoid creating an overly bureaucratic process to oversee the referee reports.

A second response might be to accept that RECs may find it difficult to stop themselves from considerations of the science, for the reasons we outline above. If RECs are to consider such matters, then their membership needs to include the appropriate expertise to provide the knowledge and credibility necessary to the (legitimate) exercise of power. To avoid the outcome of an application being (overly) dependent on the particular compositions of individual committees, committees would also need to be constituted so that their expertise was commensurate with the applications being submitted. Indeed the present situation, in which RECs are encouraged to include a statistician among their membership, goes some way towards acknowledging this implicitly. An argument might more strongly be made that committees should be constituted explicitly with methodological expertise in particular domains—clinical trials, qualitative research and so on—and only review applications within those areas. However, given the increase in multidisciplinary, multi-method studies, some flexibility in this approach would be needed.

None of these remedies is likely to completely resolve the problems (at least from the point of view of researchers) of RECs functioning both as scientific authorities as well as moral authorities in having a role in assessing both the scientific and the moral credentials of researchers and research proposals. For the present, researchers intending to conduct investigations in healthcare should recognise the degree to which their proposed research is likely to attract scientific scrutiny from RECs, and debates in the area should recognise that there may be more to this scrutiny than simply “ethics creep”.

Acknowledgements: We gratefully acknowledge funding from the National Research Ethics Services (formerly the Central Office for Research Ethics Committees) for this work, although the views expressed are the authors’ own.

Competing interests: None declared.

Ethics approval: Our study was deemed by COREC not to require REC review.

This paper was written while M D-W was a Distinguished Visiting Fellow at Queen Mary, University of London.

REFERENCES
1. Dingwall R. An exercise in fatuity: research governance and theemascula-
40:685–95.

2. The Academy of Medical Sciences. Personal data for public good: using health
information in medical research. London: Academy of Medical Sciences, 2006
Available from: http://www.acmedsci.ac.uk/images/publication/Personal.pdf
(accessed 20 November 2007)

3. Hannigan B, Allen D. A tale of two studies: research governance issues arising
from two ethnographic investigations into the organisation of health and social care.
Int J Nurs Stud 2003;40:685–95.
Organisational matters

4. Lux AL, Edwards SW, Osborne JP. Responses of local research ethics committees to a study with approval from a multicentre research ethics committee. BMJ 2000;320:1182–3.
5. Edwards SJL, Ashcroft RE, Kirchin S. Research ethics committees: differences and moral judgement. Bioethics 2004;18:408–27.
6. Haggerty KD. Ethics creep: governing social science research in the name of ethics. Qual Sociol 2004;27:391–414.
7. Diener E, Crandall R. Ethics in social and behavioral research. Chicago: University of Chicago Press, 1978.
8. Department of Health. Research governance framework for health and social care. London: Department of Health, 2001. http://www.dh.gov.uk/assetRoot/04/01/47/57/04014757.pdf (accessed 20 November 2007).
9. Central Office for Research Ethics Committees. Governance Arrangements for Research Ethics Committees (GAREC). London: Department of Health, 2001.
10. Department of Health. Report of the ad hoc advisory group on the operation of NHS research ethics committees. London: Department of Health, 2005 http://www.dh.gov.uk/assetRoot/04/11/24/17/04112417.pdf. (accessed 20 November 2007).
11. Dawson AJ, Yentis SM. Contesting the science/ethics distinction in the review of clinical research. J Med Ethics 2007;33:165–7.
12. Altheide DL. Ethnographic content analysis. In: Lewis Beck MS, Bryman A, Liao TF, ed. The Sage encyclopaedia of social science research methods. Thousand Oaks, CA: Sage, 2004.
13. Altheide DL. Qualitative media analysis. Thousand Oaks, CA: Sage, 1996.
14. Angell EL, Jackson CJ, Ashcroft RE, et al. Is “inconsistency” in research ethics committee decision-making really a problem? An empirical investigation and reflection. Clin Ethics 2007;2:92–8.
15. Evans I, Thornton H, Chalmers I. Testing treatments: better research for better healthcare. London: British Library, 2006.
16. Gieryn T. Boundary work and the demarcation of science from non-science: strains and interests in professional ideologies of scientists. Am Sociol Rev 1983;48:781–95.

Keep up to date: sign up for our alerting services

Find out automatically when an article is published on a specific topic or by a particular author. We can also alert you when an article is cited or if an eLetter or correction is published. You can also choose to be alerted when a new issue is published online [and when we post articles Online First]. Check out the New Content Alerts and Citation tracker from the Online tools section on the home page.