Ethics in action: Approving and improving medical research with human subjects

de Jong, Jean Philippe

Citation for published version (APA):
COMPLIANCE OR QUALITY IMPROVEMENT?¹

Research monitoring by US medical institutions to protect human subjects

Jean Philippe de Jong, Myra C.B. van Zwieten, Dick L. Willems

Abstract

In recent years, to protect the rights and welfare of human subjects, institutions in the United States have begun to set up programmes to monitor ongoing medical research. These programmes provide routine, on-site oversight, and thus go beyond existing oversight such as investigating suspected misconduct or reviewing paperwork provided by investigators. However, because of a lack of guidelines and evidence, institutions have had little guidance in setting up their programmes. To help institutions make the right choices, we used interviews and document analysis to study how and why eleven United States institutions have set up their monitoring programmes. Although these programmes varied considerably, we were able to distinguish two general types. ‘Compliance’ programmes on the one hand were part of the institutional review board (IRB) office and set up to ensure compliance with regulations. Investigators’ participation was mandatory. Monitors focused on documentation. Investigators could be disciplined, and could be obliged to take corrective actions. ‘Quality-improvement’ programmes on the other hand were part of a separate office. Investigators requested to be monitored. Monitors focused more on actual research conduct. Investigators and other parties received feedback on how to improve the research process. Although both

types of programmes have their drawbacks and advantages, we argue that if institutions want to set up monitoring programmes, quality improvement is the better choice: it can help foster an atmosphere of trust between investigators and IRB, and can help raise the standards for the protection of human subjects.

### 3.1 Introduction

Many authors have argued that ethical review of medical research proposals is not enough to protect the rights and welfare of human subjects, and that independent oversight on the actual conduct of research is required because it can: protect the rights and welfare of subjects; help educate investigators about ethical research conduct; improve the ethical review process; protect institutions from governmental action; and help maintain public confidence in medical research and research ethics committees (RECs) (Berry 1997, Bortolussi and Nicholson 2002, Christakis 1988, Douglass et al. 1998, Heath 1979, Hoffman 2001, Linden-Laufer 1997, McCusker et al. 2001, Norton and Wilson 2008, Smith et al. 1997, Weijer et al. 1995). Following these authors, we call this type of oversight ‘monitoring’ and define it as independent, routine, on-site oversight for human subjects protection. So, monitors are independent from the investigative team, which contrasts with data and safety-monitoring boards; monitoring is a routine activity, which contrasts with investigations of suspected misconduct (for-cause audits); and monitors go on site and contact the investigative team and the project, which contrasts with reviewing paperwork prepared by investigators for yearly re-approval (Heath 1979).

United States’ federal regulations echo this requirement for monitoring and state: ‘An IRB [institutional review board; the US terminology for RECs] shall conduct continuing review of research ... and shall have authority to observe or have a third party observe the consent process and the research’ (Department of Health and Human Services 1997). As US IRBs are part of institutions, this has been interpreted as granting both the institution and the IRB the authority to monitor (Prentice et al. 2006). In line with this requirement, many US institutions have developed monitoring programmes (VandenBosch and Maio 2011). However, there are no guidelines for the frequency,
scope, or operational methods for monitoring. Furthermore, scientific evidence on US monitoring programmes is limited to a brief 2008 report on the programmes at eleven institutions (Bortolussi and Nicholson 2002, VandenBosch and Maio 2011). Evidence from other countries is also scarce. So, institutions have practically no guidance for setting up monitoring programs, which might result in programmes that fail to protect the rights and welfare of human subjects.

To help institutions make the right choices when considering how to set up their monitoring program, we have investigated how US institutions carry out this type of oversight. In this article we give a detailed, systematic overview of the various ways institutions have set up monitoring programmes, and discuss the considerations underlying these programmes.

3.2 Method

In January and February 2008, JJ interviewed IRB members/chairs and other professionals involved in monitoring programmes at eleven US institutions (Table 1), using a structured interview format (Table 2).

To create a mixed dataset we chose six institutions with a medical school and five smaller, non-academic medical centres. The number of active, IRB-approved studies per institution ranged from 630 to 2,200. For practical reasons we focused on the north-eastern US. We also analysed documents on monitoring programmes available on the institutions' websites. These documents were first searched on the basis of the interviews in 2008, followed by systematic extraction of all relevant documents on 18/19 April 2011. Furthermore, five American experts on research ethics and IRBs (who had also served on IRBs in the past) were interviewed to ensure that our sample of institutions did not miss important aspects of monitoring practices in the US. All interviews were transcribed verbatim.

Following the structure of our interview guide, our material was coded inductively by comparing the material from different institutions, and codes were readjusted to obtain an optimal fit between the codes and the material. (Strauss and
Table 1: Study population: eleven US institutions

- New York University School of Medicine
- Johns Hopkins Medicine
- Columbia University Medical Center
- New York State Psychiatric Institute
- Albert Einstein College of Medicine
- Yale University School of Medicine
- Montefiore Medical Center
- Joan and Sanford I. Weill Medical College of Cornell University
- Partners Healthcare System for Brigham and Women’s Hospital, Faulkner Hospital, and Massachusetts General Hospital
- Saint Vincent Catholic Medical Centers
- Children’s Hospital Boston, Harvard Medical School

Table 2: Structured interview format

At this particular institution:

- What is the position of monitors within the organization?
- What is the objective of monitoring?
- How are studies selected for monitoring?
- How does the process of monitoring work?
- What additional tasks do monitors have?
- What are the effects of monitoring?

of monitoring activities and underlying considerations. Hereafter, individual institutions were used as units of analysis to discern different types of monitoring programs. This was inspired by a distinction made by several interviewees between a ‘compliance’ philosophy and a ‘quality-improvement’ philosophy. We decided to apply this distinction systematically, and named the corresponding types of monitoring ‘compliance monitoring’ and ‘quality-improvement monitoring’.

Following our interview guide, we describe the range of monitoring activities in six results sections. Within each section we report whether results apply to compliance monitoring, quality improvement monitoring, or both.

3.3 Results

3.3.1 What is the position of monitors within the organization?
Because one institution did not monitor at all, our results on actual monitoring activities are based on ten institutions.

Monitors were part of the IRB office in eight institutions, and part of an independent office in two institutions. Monitoring programmes in the latter group clearly followed a quality-improvement philosophy, and an interviewee explained why the monitors were independent from the IRB:

'We want to … maintain… the confidentiality and trust of the investigator … Investigators talk very openly and freely to [the monitor].'

These quality-improvement monitors were called quality-improvement staff, or simply investigators. Monitors who were part of the IRB office followed more of a compliance philosophy, and were called compliance monitors or auditors. Monitoring was often called an audit. An interviewee explained that a shift from compliance monitoring towards quality-improvement monitoring was reflected by changing names:

'We’re moving away from the term “audit”. We’re calling it “review”, because it’s much friendlier … “Audit” is very bureaucratic, very governmental … Our goal … is to educate them, not to find what they’ve done wrong.'
3.3.2 What is the objective of monitoring?
According to all interviewees, monitoring should protect the rights and welfare of subjects. However, compliance programmes and quality-improvement programmes translated this into different practical objectives. As one interviewee put it, the objective of compliance monitoring was the ‘enforcement of regulatory compliance’. Furthermore, some compliance programmes were initiated after incidents with human subjects and problems with governmental agencies: ‘In the past we found [that investigators] … did not do well on FDA inspections.’ According to this interviewee, monitoring should also protect the institution from governmental interference and legal action:

‘An institution could be held liable if something goes wrong … [If an investigator is not in compliance] an institution is exposed without any [insurance] coverage.’

Quality-improvement programmes framed the goal of subject protection in terms of quality improvement: ‘[to] continually … improve the research process.’ Furthermore, according to our interviewees, quality-improvement monitoring ‘is a service [to investigators]’ and helps ‘to understand where, as an institution, we may be able to improve.’ One interviewee explained why they had chosen a quality-improvement philosophy rather than a compliance philosophy:

‘The culture around here [is] about … openness … Communication and trust just permeates everything [here] with clinical research … I don’t want to disassemble … [this] culture … [That’s] why [our monitors] don’t … go directly to the IRB [with their findings].’

3.3.3 How are studies selected for monitoring?
Four monitoring programmes selected studies at random, four used a risk-based approach, and two used a combined approach. Table 3 summarizes the criteria for selecting studies in risk-based programmes. Selection methods were not related to the type of program. The number of studies monitored ranged from six to over a hundred per year, and the percentage ranged from a few percent to more than 90 percent.
Table 3: Criteria for selecting studies in risk-based monitoring programmes

Study-related risk factors:
- Studies for which an employee of the institution holds the Investigational New Drug or Investigational Device Exemption (and therefore has to fulfil associated FDA regulations)
- High safety risks to subjects (for example, ‘research that involves recombinant DNA, infectious agents and/or pathogens, biological toxins, or gene transfer or pathogens introduced into human participants’)
- Complicated study protocols
- Old research projects (for example, ‘more than five years’)
- Investigator-initiated research with no external sponsors

Investigator-related risk factors:
- Studies in which the principal investigator has recently changed
- Investigators who have come to the attention of the monitor through ‘prior experience with the responsible investigator and research team’
- Investigators who seem to be making mistakes
- ‘Gut feelings’
- Investigators new to the institution
- Inexperienced investigators
- Investigators who ‘seem slick’

Report-related risk factors:
- Problems with continuing review applications and progress reports
- Inadequate adverse-event reporting

In both types of programmes, studies were also monitored to help ‘prepare for upcoming [external] audits by the FDA, NIH [National Institutes of Health], [or] sponsors.’ Quality-improvement programmes also monitored at the request of investigators and in one programme this was the standard approach; as one interviewee said, ‘We don’t contact people anymore, they call us up: “Would you please come?”’
3.3.4 How does the process of monitoring work?
All monitors sent investigators a letter to announce a monitoring visit. Some monitors specified what information they would need, allowing investigators to prepare for the visit. In quality-improvement programmes visits were voluntary, as phrased on a website: ‘The selected PI [principal investigator] has the option to defer or decline participation,’ whereas in compliance programmes visits were mandatory.

Monitors usually started visits by meeting briefly with the investigator. Visits took from two hours to two days, depending on the complexity of a study and the type of programme: compliance programmes generally taking longer. An interviewee explained why quality-improvement visits were relatively short:

‘If you go there and live with them for a week, nobody wants you to come … [Furthermore] I trust … [investigators]. I don’t want … [them] to show me everything … [but] some things [that indicate they] … are in good shape. We don’t want to oppress [them].’

All monitors checked whether study records and documents (Table 4) complied with good clinical practice guidelines and regulatory requirements for the protection of human subjects, and corresponded with the IRB’s records.

Quality-improvement monitors also toured the research facility and talked with the research staff. As one interviewee explained:

‘We’ll ask them: … Where are your barriers? What are the problems? … What’s the informed consent process? … That’s the way we get feedback.’

Interviewees of both types of programmes reported that on rare occasions (if the research presented significant risks or if subjects would probably have difficulty understanding relevant information) the informed consent process would be observed.

Final reports were sent to investigators and, in compliance-monitoring programmes, also to other institutional officials. Quality-improvement reports remained confidential unless monitors found ‘serious and continuing non-compliance’, reportable to the authorities and reason to start a for-cause audit. According to a website, quality-improvement reports were made ‘for quality-improvement and educational purposes’ and contained voluntary recommendations to ‘help … improve identified “problem”’
### Table 4: Documents checked by monitors

**Basic regulatory documents:**
- Protocol (original and amended versions)
- Consent forms with different versions
- The research team’s curricula vitae, licenses, and certifications
- Laboratory certifications and normal values
- Staff signature log
- Delegation of responsibility log
- Enrolment log
- Documentation of protocol deviations
- Documentation of data- and safety-monitoring activities
- Documentation and assessment of adverse events
- Correspondence with the study sponsor: safety reports, monitoring reports, adverse-event reporting, correspondence on protocol deviations
- Correspondence with the IRB: submitted study documents, approvals of study documents, reports of adverse events, reports of protocol deviations, reports of unanticipated problems

**Additional regulatory documents for investigational drug or device studies:**
- Investigator’s brochure
- Drug/device accountability log
- Correspondence with the FDA

**Individual subject records:**
- Informed consent forms: version, signatures, dates
- Inclusion/exclusion checklists
- Research data
- Source documentation verifying eligibility and research data

areas’. This contrasts with compliance reports that contained mandatory corrective actions, as is illustrated by this quote:
The report will include any findings that need to be addressed, specified corrective actions, and a time frame by which the action needs to be completed.

Several monitors of both types would follow up on visits by requesting status reports or conducting a second visit. Again, in quality-improvement programmes this was voluntary, whereas in compliance programmes this could be a 'mandatory reassessment at a future date'. Furthermore, in compliance programmes, the IRB could take punitive actions against investigators. IRBs did not have 'black-and-white, hard rules' to guide such actions, but took several factors into account, as one compliance monitor explained:

'[Were] the subjects really at risk? … Was anybody harmed? … Was it done wilfully? … Was the PI … acting responsibly … or … just left it up to … people who didn’t know what they were doing? … [Has] that researcher … been in trouble with [the IRB] … before?'

IRBs linked to compliance programmes could take various actions: demand good clinical practice training, issue a warning, (temporarily) stop the study, restrict the investigator's ability to do research, fire the investigator, or, as is illustrated by this quote:

'If [the investigator] … submit[s] … another research protocol [and has] … a bad history with us … [the IRB is] going to be harder on him.'

Quality-improvement programmes however, did not use punitive action, as one interviewee explained:

'We [the monitors] don’t have the authority to do anything, but that’s why … investigators like us … We make recommendations.'

3.3.5 What additional tasks do monitors have?

All but one monitors (a compliance monitor) undertook activities besides monitoring.

Both types of monitors conducted for-cause audits at the request of the IRB or other institutional officials in order to respond to complaints, reports or suspicions of non-compliance (including monitoring reports), or concerns from governmental agencies.
Both types of monitors also had educational tasks, as one interviewee explained: 'Communicating ... information to investigators ... is ... [the monitors'] number-one priority.' Monitors educated through participation in trainings, presentations and symposiums, providing information through newsletters and websites, providing tools for organizing study documentation and providing assistance with study start-up. Although the kind of educational activities overlapped in compliance and quality-improvement programmes, the latter included more educational activities, and focused less on regulations but more on practical advice on running a well-organized study.

Quality-improvement monitors also used the information collected during visits to give the IRB feedback on the review process and institutional officials 'feedback ... about what burdens, what barriers [investigators] have in doing ... research' in order to improve institutional policies and research facilities. An interviewee explained:

'We look at it [as] ... quality-improvement of the whole, any part of the institution that has anything to do with human research protection. When we review an investigator, at the same time we also review the IRB ... [their] record [and] ... decision.'

Furthermore, feedback was used to improve the quality-improvement monitoring programmes, as another monitor explained:

'We always ... collect information that improves our tools, [to make them] not only efficient for us to use, but also user-friendly for study sites.'

3.3.6 What are the effects of monitoring?

In general, interviewees of both types of programmes reported that monitors seldom found problems concerning subjects' rights and welfare, as one interviewee put it:

'Big issues have not been brought to my attention. We're not out there killing patients.'

Although all interviewees believed that monitoring had improved study documentation, most were unsure whether monitoring had improved subjects' rights and welfare: only one interviewee said that compliance monitoring had improved adherence to inclusion and exclusion criteria, and another interviewee said their quality-improvement programme had improved the informed consent process:
'A lot of people worry a lot about the form, and ... we've tried ... to ... educate them about how important ... the process [is] of how you talk to people when you approach them.'

However, only one monitor, of a quality-improvement programme, had actually measured the quality of studies before and after monitoring visits to substantiate claimed positive effects of monitoring, and found that study documentation had improved.

Some compliance monitors thought that monitoring had disturbed the relationship between the IRB and investigators, as one interviewee phrased it: 'Some researchers may feel that [the IRB is] ... policing them.' Negative effects of monitoring reported by many interviewees of both types of programmes were the investment of time by investigators and the use of institutional resources. For one institution, lack of resources was reason not to monitor at all.

### 3.4 Discussion

To summarize how ten US medical institutions monitored human subjects research: although monitoring programmes differed in many respects, they gravitated towards two general types – ‘compliance monitoring’ and ‘quality-improvement monitoring’.

Compliance monitors were part of an IRB office, and visits were called ‘audits’. The objective was to ensure compliance with regulations. Investigators could not decline participation. Monitors focused on documentation. Reports of visits were sent to the IRB, which could request mandatory corrective actions or take punitive measures.

Quality-improvement monitors were part of an independent office, and visits were called ‘reviews’. The objective was to improve the research process. Investigators requested visits. Monitors focused more on actual research conduct. Confidential reports of visits provided voluntary recommendations, and were used to improve IRB review and institutional research policies and facilities.

Although several authors have reported on monitoring activities that were part of research projects, only VandenBosch and Miao have described regular monitoring programmes at US institutions (2011). Furthermore, monitoring activities described in
three Canadian studies appear to be similar to the activities described in this study, except McCusker et al., who reported that research subjects were routinely interviewed (Lavery et al. 2004, McCusker et al. 2001, Norton and Wilson 2008). Our finding that monitoring programmes vary between institutions is also in line with the study by VandenBosch and Miao (2011). Our distinction between a compliance philosophy and a quality-improvement philosophy is supported by other authors on human subject protection at medical institutions (De Jong 2011, Koski 2003, Noah 2004, Weijer et al. 1995). Moreover, we believe this distinction fits a much broader distinction between what sociologists call ‘civic’ and ‘industrial’ worlds (Boltanski and Thévenot 2006). Our finding that institutions use a compliance philosophy or a quality-improvement philosophy to monitor research can thus be explained by the fact that institutions have both a clear social responsibility and are part of the research industry.

Our study has two limitations. Since our findings reflect what interviewees reported about programmes, observing monitoring activities in practice might have revealed additional or incongruous information. Furthermore, although we used a mixed dataset, and our findings were supported by interviews with scholars not personally involved in the monitoring programmes in this study, the validity of our conclusions rests on a relatively small dataset.

Although it is possible that institutions in other parts of the US and institutions without federal funding use different types of monitoring, we have no reason to believe that our findings do not apply to many other US institutions. Furthermore, our findings might be relevant to institutions in other countries which require monitoring: European Union guidelines require that sponsors monitor the conduct of clinical trials, which implies that if an institution is the sponsors, e.g. investigator-initiated studies, the institution should monitor; (The Commission of the European Communities 2005) and regulations in Canada, Australia, and New Zealand require that RECs monitor (Douglass et al. 1998, McNeill et al. 1990, Weijer et al. 1995). Because the practice of medical ethical review is quite similar in the US, the United Kingdom, Canada, Australia, and New Zealand (Fitzgerald et al. 2006), we suggest that monitoring programmes could also be similar.

Our findings provide institutions with an overview of the possible ways of setting up a programme in order to meet the ethical and regulatory requirement to
monitor medical research. Furthermore, our findings highlight two important considerations for deciding how to set up a programme.

First, institutions have to consider to what extent they want to monitor. Our results lend some support to two reasons to monitor, discussed in the literature (Berry 1997, Bortolussi and Nicholson 2002, Christakis 1988, Douglass et al. 1998, Fitzgerald et al. 2006, Heath 1979, Hoffman 2001, Linden-Lauffer 1997, McCusker et al. 2001, Norton and Wilson 2008, Smith et al. 1997, Weijer et al. 1995). The first reason is that monitoring protects the rights and welfare of subjects. Most interviewees in our study were not sure whether monitoring had improved subjects’ protection. However, monitors of both types of programmes thought that monitoring had improved study documentation. We think it is not unreasonable to expect that carefully executed studies have well-organized study documentation, and therefore that study documentation might function as a ‘surrogate outcome measure’ for human subjects protection. However, we also think that the relationship between study documentation and human subjects protection, and generally spoken, finding the most effective approach for protecting human subjects, warrants further investigation in order to prevent monitoring from becoming merely an additional layer of bureaucracy. The second reason to monitor is that it helps educate investigators about ethical research conduct. We found, indeed, that monitors considered education an important task. Again, although it is plausible that this contributes to human subjects protection, it was unknown whether this was the case.

Our results also support an important reason against extensive monitoring: the investment of time by investigators and use of institutional resources. We therefore suggest that in addition to considering effectiveness, institutions also take cost-effectiveness into account when deciding to what extent they will monitor. Our study shows one important way to maximize cost-effectiveness: select studies based on risk factors (table 3), so that only those studies are monitored for which subjects’ protection is needed most.

The second important consideration for setting up a monitoring programme which is highlighted by our findings, concerns the kind of programme. We have shown that institutions deploy a broad range of monitoring activities, which can be explained by the lack of guidance for institutions (VandenBosch and Maio 2011). In our opinion, a
lack of detailed official guidance is not necessarily problematic, since it allows institutions to set up programmes that suit their specific needs, organizational structure, and research culture. We have also shown that there is a more fundamental choice to be made between embracing a compliance philosophy or a quality-improvement philosophy. Our results indicate that both types of programmes have specific drawbacks and advantages in addition to the more general reasons to monitor or not, described above.

The strong focus on documentation issues appears to be both an advantage and a drawback of compliance programmes. According to our interviewees, an advantage of compliance programmes is that enforcing regulatory requirements concerning study documentation (Table 4) helps to protect the institution from federal (e.g. FDA) involvement and legal liability. However, interviewees also acknowledge that it is unclear whether focusing on documentation is a good way to protect human subjects. Since quality-improvement programmes use additional methods (e.g. touring the research facility, talking to research staff) to monitor, they have more opportunities to offer such protection. So, the focus on documentation is also a limitation of the compliance approach.

Another drawback of compliance programmes is related to the fact that they are carried out by the IRB office and can amount to punitive measures. According to our interviewees, this can damage the atmosphere of trust between investigators and IRB. We believe that this, in turn, could destroy the ‘informal monitoring system’, i.e. people voluntarily helping the IRB with research oversight by identifying problems related to subjects’ protection (Levine 1980). Because quality-improvement monitoring is not carried out by the IRB and findings are not directly communicated to the IRB (except in cases of serious research misconduct), it can help foster the trust fundamental to the existence of the academic community (Levine 1980). Contrary to our findings, we suggest that in order to remain trusted by investigators, quality-improvement monitors should not (further) investigate suspected research misconduct affecting human subjects (i.e. for-cause audits). Furthermore, although some people believe that increased oversight on research conduct will increase public trust, this is not necessarily true for oversight that is perceived as policing, which is, according to our study, the case with compliance monitoring (Holmberg 2004, Weijer et al. 1995).
A specific advantage of the quality improvement programmes we studied is that although the findings of visits remain confidential, they are also used, in a non-identifiable way, as feedback: to the IRB in order to improve the review process, e.g. feedback that the IRB needs to communicate the regulatory background of their demands; and to the institution in order to improve research policies and facilities, e.g. feedback concerning inadequate study support staff. We believe that this helps to foster the atmosphere of trust between investigators and the IRB, and helps the institution to raise its standards for human subjects protection.

All in all, we believe that our study indicates that monitoring according to a quality-improvement philosophy is better suited to protecting human subjects than monitoring according to a compliance philosophy. We therefore recommend that if an institution wants to set up a monitoring program, this should be done according to a quality-improvement philosophy.
References


