

Effectiveness of a responsible conduct of research course: a preliminary study

Sean T. Powell · Matthew A. Allison · Michael W. Kalichman

Received: 20 August 2005 / Accepted: 11 May 2007 / Published online: 19 June 2007
© Springer Science+Business Media B.V. 2007

Abstract Training in the responsible conduct of research (RCR) is required for many research trainees nationwide, but little is known about its effectiveness. For a preliminary assessment of the effectiveness of a short-term course in RCR, medical students participating in an NIH-funded summer research program at the University of California, San Diego (UCSD) were surveyed using an instrument developed through focus group discussions. In the summer of 2003, surveys were administered before and after a short-term RCR course, as well as to alumni of the courses given in the summers of 2002 and 2001. Survey responses were analyzed in the areas of knowledge, ethical decision-making skills, attitudes about responsible conduct of research, and frequency of discussions about RCR outside of class. The only statistically significant improvement associated with the course was an increase in knowledge, while there was a non-significant tendency toward improvements in ethical decision-making skills and attitudes about the importance of RCR training. The nominal impact of a short-term training course should not be surprising, but it does raise the possibility that other options for delivering information only, such as an Internet-based tutorial, might be considered as comparable alternatives when longer courses are not possible.

Keywords Attitudes · Behavior · Knowledge · Research ethics · Skills

Introduction

The teaching of the responsible conduct of research (RCR) is now commonplace in many research institutions. This was not always the case, but it has been encouraged

S. T. Powell · M. A. Allison · M. W. Kalichman (✉)
Research Ethics Program, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA
92093-0612, USA
e-mail: kalichman@ucsd.edu

both by a perceived need [1] and by requirements to provide such training [2]. Surprisingly, little is currently known about the success of RCR education programs in achieving any specified outcomes.

The purposes of RCR education are not immediately apparent. In general, the goals for any kind of teaching can be divided into four broad categories: (1) knowledge; (2) skills; (3) attitudes; and (4) behaviors. Knowledge, skills, and attitudes correspond directly with the cognitive, psychomotor, and affective domains defined by Bloom [3], while the pedagogical hope is that the end result will be a change in behavior. Corresponding examples of learning objectives for RCR education could include to: (1) be able to name the principles outlined in the Belmont report, (2) be able to demonstrate moral reasoning skills, (3) demonstrate an attitude that discussions on the responsible conduct of research are worthwhile, and (4) give credit where credit is due. Further, the wide range of credible outcomes might be short- or long-term, for individuals or a community, of greater or lesser importance to the responsible practice of science, and measurable or not. Unfortunately, the evidence that RCR education meets even a small part of the range of possible goals is scarce [4–7].

The best evidence for positive outcomes of RCR education is for the skill of moral reasoning [8, 9]. Although this is not always the case [10], it is worth considering that moral reasoning skills alone are insufficient to ensure responsible conduct. Furthermore, it is arguable that even some of the most egregious cases of research misconduct were committed by individuals who had the capacity for excellent moral reasoning, but lacked the attitude or will necessary to apply those skills. Similarly, insufficient knowledge of relevant facts or resources could mean that excellent moral reasoning will still result in a flawed outcome.

The present study was designed to assess the impact of a very short-term RCR education experience on selected outcomes that might be classified as examples of knowledge, skills, attitudes, and behavior. Importantly, this is not a study to assess whether all RCR courses are effective. Rather, it is an initial study of only one course. This is also not a study to create an ideal RCR course, nor is it designed to prove the effectiveness of RCR education. Instead, the goal was to take a first step in looking at the effectiveness of existing RCR education programs. This approach was taken for two purposes. First, many RCR education experiences are necessarily short because they must be inserted into research experiences that are intended to be only part-time and/or to last for a short period of time. Therefore, it is worth asking if a brief experience (<6 h) in RCR education has a positive impact on any of the intended outcomes. Second, it is hoped that such a study will provide a useful starting point for the design of future studies by identifying potential areas and approaches that might warrant consideration.

Methods

Subjects

Between July and August 2003, students at the University of California, San Diego (UCSD) School of Medicine were asked to complete a survey. Candidates for this

study were medical students participating in an NIH-sponsored Summer Research Program. As part of this training program, students were required to attend a series of four training seminars, two of which were focused on the responsible conduct of research (RCR) and two of which dealt with policies and procedures for the use of animal and human subjects. The study population consisted of three groups, all of which participated in the NIH Summer Research Program between their first and second years of medical school: (1) Summer 2001; (2) Summer 2002; and (3) Summer 2003. To assess the possible influence of the RCR course on the outcome measures, surveys were administered both before and after completion of the two RCR seminars for the Summer 2003 group. This study was approved by the UCSD Institutional Review Board (#030783SX).

RCR course

The Summer RCR course consisted of four training sessions. Each session was conducted using a lecture/discussion format and was up to 1–1/2 h in duration. The focus for two of the sessions was institutional requirements for the review of research involving animal and human subjects, respectively. The remaining two sessions had a more general focus on research ethics and RCR. Handout materials and a PowerPoint slide presentation for the RCR lectures can be found at: <http://www.ethics.ucsd.edu/effectiveness>.

Survey instrument

Surveys were developed based on preliminary findings from other ongoing studies [11] and focus group discussions. Based on interviews of over 50 teachers of RCR courses, it was found that goals for instruction vary widely. However, using an iterative classification process, Kalichman and his colleagues found that these diverse goals could be broadly classified into one of four categories: knowledge, skills, attitudes, and behavior. Representative examples of those goals were selected and preliminary questions were developed to address each of these categories. An initial version of this survey was presented and refined through meetings with two consecutive focus groups of five and four students, respectively. All were student researchers and all but two were medical students from the target population. At the beginning of each focus group discussion, students were asked to complete a version of the survey containing all potential questions. After completion, the group engaged in a guided discussion to assess the clarity and appropriateness of the questions. Participants were encouraged to provide suggestions for improving the survey.

The final survey was designed to analyze the effect of the RCR training on knowledge, skills, attitudes, and behavior. To assess RCR knowledge, the survey included 12 multiple choice questions. The correct answers to these empirical questions were determined in advance by the investigators, and correct understanding of the questions was verified through the focus group discussions. To assess attitudes and behavior, seven questions asked for responses using a five-point Likert scale and two questions asked about conversations with colleagues regarding RCR.

It was not assumed that these questions should have a “correct” answer, but it was of interest instead to know whether these attitudes or conversations differed in the pre- and post-testing of the students. Understanding of these questions had also been verified in the focus group discussions. To assess ethical reasoning skills, participants were asked to respond to a brief scenario with each student randomly assigned to one of three such scenarios. Again, it was not assumed that respondents should come up with the “correct” answer, but it was assumed that one measure of ethical decision-making skills is the extent to which answers were based on recognizing what interests are at stake (e.g., the interests of individuals or of the institution). When repeat surveys were distributed to students following the 2003 seminars, scenarios were selected so that no individual received the same scenario he/she had received with the pre-course survey. Survey questions are included as Appendix 1 and a copy of the complete survey is provided at: <http://www.ethics.ucsd.edu/effectiveness>.

Surveys were distributed with a cover letter and administered via mail, email, or in-person. For the 2001 group of 34 students, contact information was not available for two students and three of the students were excluded because of prior participation in the focus group discussions; surveys were completed by 13 trainees, giving a response rate of 45% (13/29). For the 2002 group of 36 students, three students could not be contacted because of lack of contact information; surveys were completed by 14 trainees, giving a response rate of 42% (14/33). For the 2003 group, 23 of 23 students present (100% response rate) participated in the ‘pre-course’ survey and 15 of 16 students present (94% response rate) participated in the ‘post-course’ survey.

After collection, surveys were coded for entry into the database. The knowledge (factual) questions were scored against a key prepared by the survey authors. These scores, along with student responses to each question and responses to the seven Likert questions were compiled in a database file. For scoring purposes, the investigators identified possible interests at stake for each scenario. To minimize the risk of bias, coded surveys were scored independently by the three authors. The dependent variable for scoring of ethical decision-making was the number of interests identified by the respondents.

Analysis

Statistical analysis was conducted on the entire cohort and by groups (2001, 2002, 2003 pre-course, 2003 post-course). For descriptive purposes, demographic data were presented with distribution-free measures (Table 1). Because parametric tests are typically more powerful than non-parametric tests, and because the data distributions did not deviate sufficiently from normality to overwhelm the robustness of the ANOVA and *t*-tests, statistical comparisons were made using these parametric tests (Tables 2, 3, and 4). Group differences in the mean scores for the knowledge questions were determined by ANOVA for the 2001, 2002 and post-course 2003 groups. Changes in the knowledge scores for the 2003 group taken before and after completion of the RCR course were assessed with a *t*-test. Statistical analyses were conducted using SAS version 8.2 (Cary, NC).

Table 1 Prior courses in research ethics and experience in research

	2003 (Pre)	2003 (Post)	2002	2001
Number of students	23	15	14	13
Number of students who had a previous research ethics course (%)	3 (13)	1 (6)	2 (14)	3 (23)
Range of years since taking previous research ethics course	1.5–8	0	2–3	3–5
Median years of research experience	2	2	2.5	3 or more

Data shown are for respondents in 2003 before (Pre) and after (Post) the research ethics course and for respondents who had previously taken the course in 2002 and 2001. Results listed include the number of respondents (students) in each group; the number and percent of students who had taken a previous research ethics course; the range of the number of years in which a previous course had been taken; and the median number of years of research experience for all students

Table 2 Performance on knowledge-based questions

Student Groups	Questions		
	All	Explicitly covered in course	Not covered in course ^c
2003-Pre	54.2 ± 9.4	57.1 ± 10.5	50.0 ± 30.2
2003-Post	60.6 ± 7.4 ^a	65.7 ± 7.0 ^b	36.7 ± 39.9
2002	59.8 ± 10.9	62.9 ± 12.8	57.1 ± 18.2
2001	61.2 ± 9.7	63.4 ± 11.1	69.2 ± 32.5

^a 2003: Pre vs. Post (*t* test), *p* = 0.025

^b 2003: Pre vs. Post (*t* test), *p* = 0.005

^c 2001, 2002, and 2003 (ANOVA), *p* = 0.03

Mean correct scores in percent ± SD

Results

In 2003, 65 surveys were completed by medical students participating in UCSD School of Medicine's summer research program. Respondents were grouped according to the year in which they took the course: 2001, 2002, or 2003. Students in the 2003 group were asked to complete the survey both before (Pre) and after (Post) the summer research ethics course. At the time of distribution of the Pre surveys, 23 students were present and all completed the survey. At the conclusion of the course, 16 students were present, 15 of those students completed the survey. Only ten of the 15 had also been present on the first day of the course.

No significant differences were found among the three groups of respondents with respect to prior courses in research ethics or their research experience (Table 1). No more than three students in any of the groups reported having taken a previous course in research ethics. For all groups, the median research experience was no less than 2 years.

In a comparison of 2003 Pre and Post student scores (Table 2), performance on the knowledge questions improved significantly (*p* < 0.05). However, there were no

Table 3 Student attitudes

Statement	2003 (Pre)	2003 (Post)	2002	2001
Sloppy record keeping in research should be considered an example of research misconduct ^a	3.5	3.0	3.0	3.7
When publishing data that you had previously published it is not necessary to cite the previous publication	1.6	1.7	1.7	1.7
Any research use of animal subjects is acceptable as long as the experiments do not cause unnecessary pain or suffering	2.5	2.2	2.2	2.0
If you were concerned that someone senior to you was conducting experiments on cats that might as easily be conducted with frogs, then you would be willing to raise this issue with them ^b	3.9	3.9	3.1	2.9
Someone who has witnessed misconduct has an obligation to act	4.0	4.1	4.1	4.1
If you witnessed misconduct, then you would be willing to report that misconduct	3.9	4.1	3.9	3.7
Formal training in the responsible conduct of research should be required of all researchers ^c	3.9	4.3	4.1	3.8

^a Not statistically significant for a comparison of 2003 Pre and Post groups ($p = 0.122$), but differences among the 3 groups (2001, 2002, and 2003) approached statistical significance ($p = 0.085$)

^b $p = 0.022$ for differences among the 3 groups (2001, 2002, and 2003)

^c Comparisons of 2003 pre- and post-groups ($p = 0.075$) and differences among the 3 groups (2001, 2002, and 2003) ($p = 0.068$) both approached statistical significance

Averages of responses scored using a Likert Scale (5 = strongly agree, 4 = agree, 3 = neither agree nor disagree, 2 = disagree, 1 = strongly disagree)

Table 4 Number of conversations about research ethics

	2003 (Pre)	2003 (Post)	2002	2001
With other medical students ^a	0.70	1.13	0.29	0.31
With other researchers	0.87	0.80	0.29	0.38

^a $P = 0.007$ for differences among the 3 groups (2001, 2002, and 2003)

Average number of conversations about research ethics outside of class reported by students during the preceding 3 months

significant differences among the groups of students who had completed the course in 2001, 2002, and 2003. When analyses were restricted to those knowledge questions that were covered in lecture and/or discussions, the difference was more dramatic ($p < 0.005$). For these same questions, there were no significant differences in the scores for the groups of students who had completed the course in 2001, 2002, and 2003. For survey questions not specifically covered in the course, there was no statistically significant difference between Pre and Post scores for the 2003 group; however, based on analysis of variance, scores were greater for students taking the course in 2001 and 2002 than in 2003 ($p < 0.05$).

Attitudes, as assessed by scoring of statements related to RCR, were not significantly different between the Pre and Post 2003 groups (Table 3). The one difference that approached statistical significance was in response to the statement

“Formal training in the responsible conduct of research should be required of all researchers.” The tendency was for Post 2003 students to be more likely to agree with this statement than the Pre 2003 students ($p = 0.075$). In comparing responses among students who took the course in different years, some evidence for attitudinal differences was found. Specifically, one statistically significant change was in response to the statement “If you were concerned that someone senior to you was conducting experiments on cats that might as easily be conducted with frogs, then you would be willing to raise this issue with them.” Students who had most recently completed the course (2003) reported greater agreement with this statement than those who had taken the course in 2002 and 2001 ($p = 0.022$). Furthermore, in comparisons across students who had completed the course in different years, results approached statistical significance with 2003 students being less likely to agree with the statement that “Sloppy recordkeeping in research should be considered an example of research misconduct” ($p = 0.085$) and more likely to agree with the statement that “Formal training in the responsible conduct of research should be required of all researchers” ($p = 0.068$).

In a comparison of Pre and Post students for the 2003 course, the reported number of conversations about research ethics outside of class with either other medical students or other researchers was not significantly different (Table 4). However, in a comparison across years, students who took the course most recently reported more conversations with other medical students than did the respondents who took the course in 2002 and 2001 ($p = 0.007$). A similar pattern was found for discussions of research ethics with other researchers, but it was not statistically significant.

Skills in ethical decision making were assessed by scoring student responses to selected scenarios. The differences between Pre and Post scores for 2003 students approached statistical significance ($p = 0.06$), but no differences in performance were seen over time in comparisons among the 2001, 2002, and 2003 students ($p = 0.56$). It is noteworthy that students rarely identified even a small fraction of the possible ethical interests in these cases.

Discussion

The principal findings of the present study were: (1) a small, but statistically significant, improvement in scores on questions that tested knowledge of specific facts; (2) a borderline significant improvement in ethical decision-making skills; and (3) a borderline significant increase in agreement with the statement that responsible conduct of research courses should be required.

These results are not encouraging, but it is important to be clear that this study is not a referendum on either the effectiveness or the importance of RCR education. The course studied for this project was brief and limited in scope. Nonetheless, the course has been in place for nearly 10 years to meet NIH requirements for training in the responsible conduct of research. Precisely because it is inserted into a short-term (Summer) research experience, it has by necessity always been brief. Given the context in which these courses are taught, many courses are likely to be similar to this one.

The course studied was small both in duration and in enrollment. Neither of these factors is ideal given the goals for assessing effectiveness, but if the goal is to assess the effectiveness of existing RCR education programs, then it is essential that such studies be conducted. Despite more than 15 years of an NIH recommendation that RCR education should be provided for all trainees [2], not just those funded by NIH, this training is typically limited only to NIH trainees. In a recent study, only six of 50 RCR instructors reported that their courses were required for all trainees [11]. And when these courses are taught, they are often small. In another recent study, the median enrollment in a sampling of 11 RCR courses nationwide was just 20 students, and three of the 11 courses had fewer than 11 students enrolled [12]. In theory, a larger sample could be obtained by assessing multiple courses, but the results would be confounded because the courses are taught by different instructors, in very different environments, at different times, and in many different ways. The present study is a necessary first step before creating an instrument sufficient to detect effects despite the many sources of variation in a nationwide survey of numerous courses.

Of the four educational outcome variables studied (knowledge, skills, attitudes, and behaviors), improvement in student performance on knowledge-based RCR questions was the outcome most clearly shown by statistical analysis. When comparing the 2003 Pre and Post RCR course groups, we found a statistically significant increase in student performance on questions testing general RCR knowledge. When comparing the scores on the knowledge-based questions of 2003 Post students to the 2001 and 2002 groups (both of which had completed the course in previous years), there was no significant difference. A possible interpretation of this finding is that this course produced a lasting improvement in the area of general RCR knowledge, but it is also possible that these students performed better on the knowledge-based questions because of exposure to RCR issues in the course of their research program. In addition, it is worth noting the possibility that the 2001 and 2002 students performed similarly on the test not because of their initial training, but because the content of the questionnaire was reinforced by other education and training. However, the finding that RCR courses are effective in improving general student knowledge of RCR corresponds to student perceptions about RCR courses demonstrated in previous studies [12]. It has been shown that students in RCR courses perceive that of the four outcome variables, it is in the area of knowledge of RCR issues that the courses are most effective.

The statistically significant improvement found in performance on knowledge-based questions is especially impressive when considering some qualities of the RCR course studied. The course is very short, with two of four sessions devoted to the practical dimensions of conducting research with human and animal subjects. Although these are typically included in lists of RCR topics, the focus of these sessions was primarily on procedural issues rather than ethics per se. These were supplemented by two sessions taught by one of the authors (MK) that emphasized other general RCR topics through lecture and class discussion, with the majority of the teaching in small group format. Therefore, although the total course is 6 h in duration (1.5 h per session), the RCR lectures not dealing with research subjects consist of considerably less than 3 h.

This study was intentionally designed without any attempt to match the survey to the material actually covered in the course. That is, the questions were not intended to be a final exam for this particular course but instead a survey of general RCR knowledge. Few of the questions reflected primary teaching goals for the course. This may seem counter-intuitive, but the goal was not to demonstrate the effectiveness of this particular course, but to assess the positive outcomes, if any, of a short-term course experience. Unlike many other kinds of courses, it is hoped that students will be learning much of the material outside of the course. For example, it may be that if students become more aware of the ethical dimensions of the practice of research, then they are more likely to raise questions in the research environment about a wide-range of topics—not just those that were covered in the course. The result is that even if a student was not exposed to key information about RCR in the course, she might initiate conversations about RCR and learn that information outside the course. If the course had been designed solely to match to the questions on the survey, then improvements in knowledge would be less surprising. Thus, despite the small improvement in performance on the test of knowledge, it is impressive that the result reached statistical significance. Additionally, because one of the aims of the current study was to design a survey instrument that might be generalized to evaluate other courses, the questions on the survey were designed to be applicable to the field of RCR in general rather than only applicable to the specific RCR course we surveyed. As a result, the survey instrument might be applied to other courses with different instructors, student populations, or pedagogical methods.

The effect of the RCR course on attitudes was examined using a Likert scale for student responses. When comparing 2003 Post to Pre students, no statistically significant differences were found in student attitudes as assessed by any of seven statements posed on the survey. However, responses to the statement “Formal training in the responsible conduct of research should be required of all researchers” approached statistical significance ($p = 0.075$) with a tendency for the Post group to agree with the statement more strongly than the 2003 Pre group. Although this result is of borderline significance, it suggests a trend toward attitudinal changes as an effect of RCR courses. This nominal effect is consistent with the view that RCR courses tend to be effective in increasing student knowledge, while they are perceived to be less effective in changing student attitudes [12].

Examination of the effects of the RCR course on student attitudes across the three different years revealed one statistically significant difference in student responses and one result that approached significance. Compared to the 2002 and 2001 groups, students in the 2003 post group were in stronger agreement with the statement, “If you were concerned that someone senior to you was conducting experiments on cats that might as easily be conducted with frogs, then you would be willing to raise this issue with them” ($p = 0.02$). The difference observed might be attributable to the more recent completion of the RCR course. However, it is also plausible that the observed difference is due to unrelated, but inherent, differences among classes (2003 vs. 2002 vs. 2001).

Agreement with the statement “Formal training in the responsible conduct of research should be required of all researchers,” was slightly greater, although not statistically significant ($p = 0.068$), in the 2003 Post students compared with the 2002 and 2001 students. Of note, the scores for this question were inversely related to the amount of time since completion of the course. This potential relationship may reflect an increasing emphasis on clinical rather than on research experience in the second and third years of medical education.

Questions about the number of conversations students in the course had had with other medical students, and with other researchers regarding RCR in the past 3 months, were designed to assess the effect of RCR courses on student behavior. The analysis revealed no significant difference in the number of conversations between the 2003 Pre RCR course students and the 2003 Post group. There was, however, a significant increase in time spent in discussions with other medical students ($p = 0.007$), but not other researchers, between the 2003 Post students and the 2002 and 2001 classes. This is most likely a reflection of the fact that the current students (2003) were working primarily in a research environment, while past students (2002 and 2001) were now working largely in clinical environments.

To measure the effectiveness of RCR courses on ethical reasoning, students were asked to explain why a briefly described course of action was or was not ethical. Responses were graded by how many ethical interests were identified. There was a borderline significant improvement in student scores between the 2003 Pre and Post groups ($p = 0.06$), but no difference across post groups for 2003, 2002, and 2001. These findings are consistent with a possible improvement due to the training course and with persistence of that skill over time. As for improvements in knowledge, it is plausible that ethical decision-making remains high in more senior students not because of a long-lasting effect of the training but because of other aspects of the medical education program. In any case, it should be noted that student performance in analyzing the case studies was disappointingly low in all groups.

A limitation in the analysis of educational outcomes is the variable attendance of the students in the RCR course we studied. As a result of the demanding summer schedule, there was a high risk that individual students from the study population would miss one or both of the general RCR sessions. In fact, five students in the 2003 Post group had not attended the first meeting of the course. Similar patterns of attendance had occurred in earlier years of the course. Unfortunately, records of attendance were not available to be matched against the anonymous survey responses to determine which surveys reflected student responses after attending neither, one, or both RCR sessions. It is notable that despite this source of increased variation improvements in scores were still measurable.

The present study is relevant to the design and evaluation of individual RCR courses. It is reasonable to assume that effective instruction depends on clearly defined teaching objectives, a curriculum designed to meet those objectives, and measurable outcomes to determine whether they are met. This study provides a framework for thinking about what should be done in such courses. Depending on specific goals, different course approaches, durations, and assessment tools may be appropriate. If the goal is effective training, then matching these methods to the teaching objectives of a course is a necessary first step.

The outcome categories outlined in this project were knowledge, skills, attitudes, and behavior. For a focus on *knowledge*, it is hoped that trainees will learn certain information. This includes, for example, federal regulations governing financial conflicts of interest, principles of the Belmont Report, guidelines for authorship, historical examples of misconduct, or institutional resources for further information. Such material is readily covered in lectures or even Web-based tutorials. And assessment of outcomes is easily accomplished through multiple choice or “fill in the blank” exams.

A second focus for RCR courses is *skills*, particularly the skills of moral reasoning and ethical decision-making. Skills are most likely to be acquired through practice rather than merely reading or listening to others. This practical aspect means that instruction must include opportunities to struggle with the ethical dimensions of the practice of research. This is typically accomplished by discussing cases or specific problematic situations. Such discussion depends on course formats (classroom or Web-based) that challenge students with open-ended dilemmas and the expectation that they will articulate their own perspectives as well as listen to views of others. Evaluation of this outcome is considerably more difficult than for the teaching of new knowledge. A nominal approach is to recognize that the process itself is an important advance. In that case, mere participation in the discussion is a sufficient endpoint. However, if the goal is to recognize the quality of a response, then a measurable outcome is needed. The Defining Issues Test [9, 13] is one such approach, but may not be practical or even appropriate for all instructors. An alternative, that might be more value neutral, is to seek ways to address the richness of a response (e.g., identification of ethical principles as a basis for action or inaction, or recognition and definition of the many interests that are at stake).

A focus on *attitudes* is very different than knowledge or skills. It is clearly possible through repetition or practice to learn new information or skills, but that does not guarantee a positive disposition. For RCR instruction, it is likely that what happens outside the classroom will have a greater impact on attitudes than anything the instructor might do. That said, any possible success in shifting attitudes will depend on the choice of material to teach knowledge and skills, the instructor’s passion and commitment, the place of the course in the curriculum, and the ability of the instructor to highlight the relationship of the course to other evidence of an institutional commitment to RCR. Measurement of changes in attitudes might be derived from both forced choice (e.g., “Using a scale of 1–5, how would you rank the importance of RCR education?”) or open-ended questions (“How, if at all, has your attitude toward RCR been changed by this course?”).

Finally, many instructors may have the pedagogical hope that their courses will influence future *behavior* of their trainees. Such outcomes might include an absence of certain behaviors, such as research misconduct or other misbehaviors as described by Raymond De Vries and colleagues [14]. Conversely, RCR instructors might hope that trainees will not only avoid bad behavior, but will be models of the highest standards of responsible conduct in research. Such long-term goals are important, and it is reasonable to hope that they are more likely to occur with effective training to promote knowledge, skills, and attitudes. Unfortunately, reliable measurement of such long-term outcomes would be impractical, if not impossible, for assessing the

effectiveness of an individual RCR course. On the other hand, at least one behavioral outcome may change in the short-term, and is arguably a minimal expectation if other long-term behaviors are going to improve. Specifically, evidence of increased openness, discussion, and transparency are likely to occur if trainees seek out and initiate discussions about various dimensions of RCR. Such discussions normally occur rarely [5]. Pre- and post-course surveys could address perceptions of trainees about, for example, time spent on such discussions, topics covered, or the number and standing (e.g., faculty, students, or staff) of people involved in such discussions.

In conclusion, the findings of this study were that this short-term RCR course contributed to little or no improvements in knowledge, skills, attitudes, or behavior. This is not evidence for a lack of value of RCR training, but is consistent with the modest gains to be expected from very short-term training experiences. Nonetheless, these findings are encouraging for future assessments of changes that might be present with more substantial courses.

Appendix 1. Research Ethics Survey

Note that three different versions (A, B, and C) of the survey were identical except for item number 22, which is one of three case scenarios. All three scenarios are included below.

1. In what year did you take, or will you have taken, the UCSD School of Medicine summer course in *research ethics*?
2003 2002 2001 2000
2. Have you taken any other course or received training in *research ethics*? Yes
No
If yes, how many years ago did you take the course or receive training? _____
3. How many years of *research* experience do you have?
0 1 2 3 or more
4. Who owns the data collected by a researcher at UCSD?
___ (a) the person who collected the data
___ (b) the researcher who heads the research group
___ (c) UCSD
___ (d) the agency that funded the research
___ (e) don't know
5. For federally funded research, raw data and other research records must be kept [please choose the best answer]:
___ (a) until the paper is submitted
___ (b) until the paper is accepted for publication
___ (c) until 3 years after the paper is published
___ (d) until the grant is finished
___ (e) until 3 years after the last expenditure report
___ (f) indefinitely

6. Which of the following documents best summarizes the ethical principles on which human subjects research is judged in this country?
- (a) Nuremberg Code
 - (b) Belmont Report
 - (c) Declaration of Helsinki
 - (d) none of the above
7. Which of the following are primary ethical principles by which human subjects research in this country is judged? [check all that apply]
- (a) community need
 - (b) justice
 - (c) beneficence
 - (d) quality of science
 - (e) respect for persons
8. Which of the following committees has primary responsibility for reviewing research involving human subjects?
- (a) CCC
 - (b) COI
 - (c) IACUC
 - (d) IRB
9. Which of the following item(s) is/are not sufficient alone for authorship, according to guidelines of the International Committee of Medical Journal Editors? [check all that apply]
- (a) acquisition of funding
 - (b) collection of data
 - (c) general supervision of the research group
 - (d) final approval of the version to be published
10. According to guidelines of the International Committee of Medical Journal Editors, if a potential author has drafted an article or revised it critically for important intellectual content, and has given his or her final approval of the version to be published, which of the following substantial contributions is sufficient to warrant authorship? [check all that apply]
- (a) conception and design
 - (b) acquisition of data
 - (c) analysis and interpretation of data
 - (d) acquisition of funding
11. Which of the following best characterizes the criteria for authorship in the biomedical sciences?
- (a) standards of authorship vary widely
 - (b) standards of authorship are mandated by federal regulation
 - (c) standards of authorship are covered by guidelines widely accepted by practicing scientists

- ___ (d) eligibility for authorship depends more on performing the work than on generating ideas
- ___ (e) eligibility for authorship depends more on generating ideas than on performing the work
12. Which of the following is/are included under the definition of Research Misconduct in the UCSD Integrity of Research Policy? [check all that apply]
- ___ (a) fabrication
- ___ (b) falsification
- ___ (c) plagiarism
- ___ (d) financial conflicts of interest
- ___ (e) other practices that seriously deviate from those that are commonly accepted within the scientific community

For Questions 13–19, please indicate your reaction to the statement using the scale below

- 1=strongly disagree
- 2=disagree
- 3=neither agree nor disagree
- 4=agree
- 5=strongly agree

13. Sloppy record keeping in research should be considered an example of research misconduct. ___
14. When publishing data that you had previously published it is not necessary to cite the previous publication. ___
15. Any research use of animal subjects is acceptable as long as the experiments do not cause unnecessary pain or suffering. ___
16. If you were concerned that someone senior to you was conducting experiments on cats that might as easily be conducted with frogs, then you would be willing to raise this issue with them. ___
17. Someone who has witnessed misconduct has an obligation to act. ___
18. If you witnessed misconduct, then you would be willing to report that misconduct. ___
19. Formal training in the responsible conduct of research should be required of all researchers. ___
20. During the past 3 months, how many conversations about *research ethics* have you had with other medical students outside of class?
0 1 2 3 or more
21. During the past 3 months, how many conversations about *research ethics* have you had with faculty or other researchers outside of class?
0 1 2 3 or more

Case scenario A

22. Alana is a medical student researcher in the laboratory of Prof. Hayes. Prof. Hayes has received a manuscript for review for possible publication in a biomedical journal and asks Alana to review the manuscript. Alana knows that the review process is intended to be confidential, so she asks if the journal editor has been notified of this request. Prof. Hayes says that this is not necessary. Alana asks for your advice.

(a) Is Professor Hayes' answer (that notification is not necessary) ethical?

Yes No

(b) Why or why not?

Case scenario B

22. Xiao is a medical student working on his first research project. While working late in the laboratory, he notices that Claudia, a senior postdoc, left her lab notebook out and open in a common area of the lab. As Xiao picks up the notebook to return it to her desk, he sees that the entry dated for today shows measurements that would have been obtained from a piece of lab equipment for which Xiao has been the sole user for the last 2 days. It would not be possible that Claudia had made those measurements at this time. Xiao is considering that he should say or do nothing about his observation, but he asks for your advice.

(a) Is Xiao's decision (to say or do nothing) ethical?

Yes No

(b) Why or why not?

Case scenario C

22. Eduardo is a medical student assigned to an interview study of elderly subjects with Alzheimer's disease. Although these subjects had the capacity to consent at the time of the onset of the study, some of them are now suffering from severe dementia. In Eduardo's most recent interview of one of these subjects (Molly), he found that she repeatedly punctuated her answers with the statement: "Don't ask me any more questions." However, Molly continued to answer Eduardo's questions, despite a clearly diminished capacity to understand what was being asked. On the assumption that Molly had chosen to participate at a time when she still had the ability to offer informed consent, Eduardo decided to complete the interview. However, he is now asking you for your advice about keeping Molly in the study.

(a) Was Eduardo's decision (to keep Molly in the study) ethical?

Yes No

(b) Why or why not?

References

1. National Academy of Sciences (1992). *Responsible science: Ensuring the integrity of the research process, Vol. I*. Washington, D.C.: National Academies Press.
2. National Institutes of Health, and Alcohol, Drug Abuse, and Mental Health Administration (1989). Requirement for programs on the responsible conduct of research in National Research Service Award institutional training programs. *NIH Guide for Grants and Contracts*, 18, 1.
3. Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals*. New York, NY: David McKay Company, Inc.
4. Kalichman, M. W., & Friedman, P. J. (1992). A pilot study of biomedical trainees' perceptions concerning research ethics. *Academic Medicine*, 67, 769–775.
5. Brown, S., & Kalichman, M. W. (1998) Effects of training in the responsible conduct of research: A survey of graduate students in experimental sciences. *Science and Engineering Ethics*, 4(4), 487–498.
6. Eastwood, S., Derish, P., Leash, E., & Ordway, S. (1996). Ethical issues in biomedical research: Perceptions and practices of postdoctoral research fellows responding to a survey. *Science and Engineering Ethics*, 2, 89–114.
7. Committee on Assessing Integrity in Research Environments (2002). *Integrity in scientific research: Creating an environment that promotes responsible conduct*. Washington, D.C.: Board on Health Sciences Policy and Division of Earth and Life Studies, Institute of Medicine and National Research Council of the National Academies, National Academies Press.
8. Hull, R., Wurm-Schaar, M., James-Valutis, M., & Triggle, R. & D. (1994). The Effect of a Research Ethics Course on Graduate Students' Moral Reasoning. National Academy of Sciences, Washington, D.C. July 5–6, 1994. [available online at: http://www.richard-t-hull.com/publications/effect_research_ethics.pdf] Accessed 5/11/07.
9. Bebeau, M. J., Pimple, K. D., Muskavitch, K. M. T., & Smith, D. H. (1995). *Moral reasoning in scientific research: A tool for teaching and assessment*. Bloomington, IN.: Indiana University.
10. Heitman, E., Salis, P. J., & Bulger, R. E. (2000). Teaching ethics in biomedical science: Effects on moral reasoning skills. In Steneck, N. H. & Scheetz, M. D. (Eds.). *Investigating research integrity. Proceedings of the First ORI Research Conference on Research Integrity*. pp. 195–202. [available online at: http://www.ori.hhs.gov/documents/proceedings_rri.pdf] Accessed 5/11/07.
11. Kalichman, M. W., & Plemmons, D. K. (2007) Reported goals for responsible conduct of research courses. *Academic Medicine* in press.
12. Plemmons, D., Brody, S. A., & Kalichman, M. (2006). Student perceptions of the effectiveness of education in the responsible conduct of research. *Science and Engineering Ethics*, 12, 571–582.
13. Rest, J., Narvaez, D., Bebeau, M. J., & Thoma, S. J. (1999). *Postconventional moral thinking: A Neo-Kohlbergian approach*. Florence, KY.: Lawrence Erlbaum Associates.
14. De Vries, R., Anderson, M. S., & Martinson, B. C. (2006). Normal misbehavior: Scientists talk about the ethics of research. *Journal of Empirical Research on Human Research Ethics*, 1(1), 43–50.