

Protection of Human Subjects

RISKS TO HUMAN SUBJECTS

This Human Subjects Research falls under Exemption 1 (E1) for “research conducted in established or commonly accepted educational settings, involving normal educational practices, such as (i) research on regular and special education instructional strategies, or (ii) research on the effectiveness of or the comparison among instructional techniques, curricula, or classroom management methods” (Section 46.101 of 45 CFR 46). The only human subjects in this project will be trainees in the educational settings of sessions where educational material will be presented in a typical, workshop-style educational setting (on-campus pilot testing and conference workshop) and trainees who use the online training modules we plan to develop. All interactions will therefore be in a typical educational setting, with normal educational practices, and with the purpose of gaining feedback on the frequency of use and effectiveness of the training materials developed through the grant. For this project, we plan to use pilot testing to help us evaluate and improve the training modules. We ask for and record data on the number of and basic make-up (educational level, research roles) of the pilot testers and online users. We will also ask users to complete a survey to collect their feedback on the training modules and their effectiveness.

Human Subjects Involvement, Characteristics, and Design

The involvement of human subjects in this project will be in pilot testing and evaluating the training modules we propose to develop. There are three groups of human subjects involved in this pilot testing and evaluation of training modules: (1) pilot testers who participate in the six bi-annual pilot testing sessions at Colorado State University over the project period; (2) participants in the proposed workshop at the American Society of Microbiology meeting in year 2 of the project; and (3) early users of the online book and related educational materials.

Members of the first two groups will participate in day-long, workshop-style sessions where we will present the material developed for the training modules and where the participants will work through applied examples and other additional educational materials. We will collect basic characteristics of the participants of these sessions (basic demographics, education level, research role). We will ask them to take a voluntary survey at the end of the session with questions about the usefulness, clarity, and relevance of the training materials covered in the session. Six months following the session, we will send the participants a voluntary survey with questions to help assess the long-term outcomes of the training session in terms of how much of the material was helpful to the participant in improving data reproducibility in their own research projects. The immediate surveys will be given in person, while the follow-up survey will be conducted online, with a link to the online survey emailed to the participants. We anticipate that each of these sessions will have approximately twenty participants.

The third group will consist of early users who find and try out our online book and its related training materials. The book will include embedded voluntary surveys with questions on basic characteristics of the user (demographics, education level, and research role) as well as questions about the usefulness, clarity, and relevance of the training materials. The book and its materials will also be tracked with Google Analytics and other online analytics tools to track high-level patterns in the use of the materials among early users. We anticipate there will be 50–75 early online users of the in-development online training material, who will be recruited through the planned dissemination efforts with social media and among our network of R users and biomedical researchers.

We anticipate that these pilot testing groups will have a composition that is representative of the microbiology / immunology academic research community, with a mix of males and females and a wide range of adult ages. We anticipate that most participants will be healthy, and some participants may be pregnant, reflecting the average composition of our target audience of laboratory-based biomedical researchers. The participants for the in-person pilot testing will not include prisoners, institutionalized individuals, or children; we anticipate but cannot guarantee that the early online users will not include these populations, either. Recruitment for inclusion in these groups will be based on being members of the target audience for our training modules, namely academic researchers whose research focuses on laboratory-based, rather than computational, research strategies. The sites of the in-person sessions will be Colorado State University and the site of the American Society for Microbiology annual conference in year 2 of our project. The project's principal investigator (Anderson) and one or more of the co-investigators will be in charge of planning and

conducting the in-person pilot testing sessions / workshop, including the administration of surveys to collect feedback on the training materials, as well as of developing and posting the voluntary surveys to be included in the online book.

Source Materials

The materials collected will be: (1) data on the characteristics of each group of pilot testers, including basic demographics, educational level, and research role; (2) answers to survey questions on the usefulness, clarity, and relevance of the training materials; and (3) tracking data on patterns of use of the online material (collected only for early online users). We will not store names, IP addresses, or other personal identifiers with this data. We will convert this anonymized data to an electronic form and save it on Colorado State University-owned, password-protected computers. Only members of the research team will have access to this data.

Potential Risks

The risks to the participants will be minimal—no greater than the typical risks encountered in daily life. We anticipate that some of the participants may be pregnant women, for whom the risks will also be minimal.

ADEQUACY OF PROTECTION AGAINST RISKS

Recruitment and Informed Consent

We will recruit pilot testers for the Colorado State University pilot testing sessions mainly through email announcements to relevant listservs maintained for biomedical and related research at Colorado State University. We will recruit participants for the American Society for Microbiology conference workshop through the American Society for Microbiology's workshop information and registration materials and website. We will recruit early online users through the planned dissemination efforts with social media and among our network of R users and biomedical researchers. At the beginning of each of the surveys, we will provide the information needed to make an informed choice for whether to consent to the survey information being collected and used to refine the training material. We will include a disclosure at the beginning of the online book to inform users of the book that the website includes analytic tracking, including *Google Analytics*, to help assess patterns of use for the online training materials.

Protections Against Risk

Risks to participants will be minimal. All interactions will therefore be in a typical educational setting, with normal educational practices, and with the purpose of gaining feedback on the frequency of use and effectiveness of the training materials developed through the grant. We will anonymize all collected data prior to saving it electronically, and we will store this anonymized data in university-owned, password-protected computers.

POTENTIAL BENEFITS OF THE PROPOSED RESEARCH TO HUMAN SUBJECTS AND OTHERS

The key benefits to participants will be new training and career development. The key benefits to others will be well-developed training modules to improve the reproducibility of data recording and pre-processing in biomedical research. Ultimately, this can help in ensuring biomedical research that is more rigorous and more likely to improve human health. There will not be any financial compensation to participants.

IMPORTANCE OF THE KNOWLEDGE TO BE GAINED

The key knowledge to be gained will be the feedback needed to create well-developed training modules to improve the reproducibility of data recording and pre-processing in biomedical research. Ultimately, this can help in ensuring biomedical research that is more rigorous and more likely to improve human health.

Inclusion of Women and Children

INCLUSION OF WOMEN AND MINORITIES

We anticipate that the pilot testers will include women and minorities, in a proportion reflective of the composition of our target audience of laboratory-based biomedical researchers. To ensure an open and inclusive environment for all participants, including women and minorities, we will share with participants a code of conduct for the session prior to the event, based on codes of conduct enforced at the annual R Users' Conference (e.g., https://user2018.r-project.org/code_of_conduct/).

INCLUSION OF CHILDREN

We anticipate all program participants will be 18 years old or older. Children will be excluded from the proposed pilot training, as they are not representative of our target audience for the training materials of laboratory-based biomedical researchers, and so their feedback would not be useful in helping us to refine and improve the training materials.