Center for Scientific Review
Peer Review Advisory Committee Meeting
National Institutes of Health
U.S. Department of Health and Human Services

January 23, 2006

The first 2006 meeting of the Peer Review Advisory Committee (PRAC) convened at 8:30 a.m. on Monday, January 23, 2006, at the Hyatt Regency, Bethesda, Maryland. The entire meeting was held in open session. Dr. Antonio Scarpa and Dr. Jeremy Berg presided as Co-Chairs.

Members

Jeremy Berg, Ph.D., Co-Chair
Antonio Scarpa, M.D., Ph.D., Co-Chair
Dean E. Brenner, M.D.
Faye Calhoun, Ph.D.
Joe. L. Martinez, Jr., Ph.D.
Craig J. McClain, M.D.
Edward N. Pugh, Jr., Ph.D.
Louise Ramm, Ph.D.
Anne P. Sassaman, Ph.D.
Beverly Torok-Storb, Ph.D.
Matt Winkler, Ph.D.

Ad hoc Members

Leslie Leinwand, Ph.D.
Daria Mochly-Rosen, Ph.D.

Dr. Michael R. Martin was the Executive Secretary for the meeting.

Welcome, Introductions, and Approval of the September 2005 PRAC Meeting Minutes

Dr. Berg welcomed participants to the Peer Review Advisory Committee meeting and asked for approval of the minutes from the last meeting. Dr. Louise Ramm requested a correction to delete an inaccurate reference attributed to her on electronic submission of grant applications. The minutes were then approved, and the correction was later made.

Executive Secretary Michael Martin then asked the members to briefly introduce themselves. He welcomed new members: Drs. Daria Mochly-Rosen, Leslie Leinwand, and Dean Brenner.

Remarks from the CSR Director: New Challenges and Opportunities

Dr. Scarpa thanked PRAC members for attending and noted that new challenges and opportunities make their assistance particularly necessary. He said he would focus his comments on changes carried out in the past 6 months since he became the CSR Director, as well as changes still needed to deal with budget and workload challenges.

Changes in CSR Operations
Three areas of change have taken place in CSR operations: an increase in communication and transparency, an increase in uniformity, and an increase in efficiency.

Communication and Transparency: CSR staff members now spend more time sharing information with each other, as well as with the other ICs and with other agencies that carry out peer review in the United States and abroad. CSR is also engaging in communication with others in the scientific community through meetings with outside groups and Institute and Center Councils at the National Institutes of Health (NIH). He explained that he spends about half his time getting input on what needs to be done to improve peer review.

Uniformity: CSR has set standards for nominating reviewers for service on its study sections that will aid the development and approval of nomination slates. In addition, summary statements, which are CSR's measurable output, will be posted within one month of a study section meeting, with statements for new investigators' applications posted within one week, which will give applicants more time for revisions. In addition, resumes will have a standardized structure across all CSR study sections, which will benefit both NIH and the investigator. Dr. Scarpa noted that CSR has been asking its Scientific Review Administrators to unscore or streamline 50 percent of their applications, but on average only 32 percent of their applications are unscored. He said that CSR would put in place a common practice on streamlining applications.

Efficiency: By early 2007, most NIH applications will be submitted electronically, which will represent an incredible increase in efficiency. In the meantime, artificial intelligence or fingerprinting software is being considered to increase efficiency in assigning applications to particular CSR integrated review groups (IRGs) or even study sections, as well as in identifying potential reviewers. Nine pilots are underway to assess the benefits of various knowledge management tools.

Upcoming Changes in CSR Operations
New speaker phones will be used in study section meetings so that program staff can attend remotely, rather than rushing to and from meetings. The phones will be available in all meetings by February 1, 2006.

Dr. Scarpa noted the original goal when study sections were realigned was to have an external review on a 5-year cycle. However, this cycle can be too long, given the pace of scientific advances, for some issues. Thus, one IRG will be reviewed internally each month, so that a total review of all IRGs occurs every 2 years. This will supplement the 5-year external review cycle.

Desirable Changes in CSR Review
Stakeholders have expressed concerns about the long review cycle, whether clinical research is properly evaluated, and improved assessment of innovative, high-risk/high-reward research. Recruiting and retaining more high-quality reviewers is another area of possible change to the current systems.

Shortened review cycle: In addition to posting summary statements sooner, Dr. Scarpa said that a pilot study is underway in 40 CSR study sections in which new investigators will have the

opportunity to revise and resubmit an application in the very next review cycle.

Clinical research: Dr. Scarpa acknowledged the concern in the community that clinical research is not properly evaluated. He explained that studies conducted on this topic will also be discussed later on the PRAC agenda, noting that the most recent data analyzed suggests that clinical research applications are not disadvantaged in reviews.

Innovative research: Different constituents say that conservative proposals score better than innovative ones. Dr. Scarpa shared a published criticism by the president of the American Society for Cell Biology.

High-quality reviewers: A review is as good as the reviewers, said Dr. Scarpa, and CSR alone uses 18,000 reviewers per year. CSR is experimenting with different platforms to recruit and retain high-quality reviewers. For example, new technology is improving telephone- and video-enhanced discussions. Asynchronous electronic discussions can involve reviewers who cannot travel or have schedules or live in time zones that prevent the entire study section from meeting at the same time. Three such reviews have taken place with physicists and computational biologists, and they went well. Asynchronous discussion is useful for international reviewers, who are valued, given the general need for reviewers and the difficulty of finding U.S. reviewers with no conflicts of interest for large grants that may involve 10 or 20 institutions.

Increase in Applications

Dr. Scarpa shared data on the increase in the number of applications. In addition to about 9,000 more applications coming from new investigators, more applications are being submitted per person. For 20 years, investigators submitted an average 1.2 applications per year. Beginning in 2002, that number jumped to 1.4 applications—which results in about 20,000 additional applications annually. At the same time as applications are increasing, reviewers want to read fewer applications each, especially reviewers in clinical departments. Meanwhile, the CSR budget has remained constant, with non-discretionary spending squeezed each year by mandatory cost increases.

Business as usual is no longer possible, and Dr. Scarpa said that the challenges exist in terms of demands on both reviewer time and CSR staff time. He listed and briefly discussed 10 possible short-term approaches to increasing efficiency for reviewers and/or CSR: (1) shifting additional grant reviews to Institutes; (2) replacing many special emphasis panels (SEPs) with smaller parallel study sections; (3) enlarging study section membership to decrease the frequency of participation; (4) increasing pre-meeting streamlining; (5) using various review platforms; (6) using hybrid review platforms; (7) staggering application deadlines; (8) assigning two rather than three reviewers to an application; (9) shortening applications; and (10) instituting more structured applications and reviews. Dr. Scarpa closed by underlining the value of peer review to NIH and to scientific advances, and noted the pride his CSR colleagues take in their work.

Discussion

Dr. Joe Martinez said that another way to limit the number of applications is to limit the number an investigator can submit. He wondered if any evidence shows that more grants produce better science. Dr. Scarpa said that his impression is that many investigators need two grants to cover

salaries. Dr. Norka Ruiz Bravo estimated that perhaps 200 investigators have three to four grants, although she did not have exact numbers. Dr. Berg noted that the National Institute of General Medical Sciences (NIGMS) has a policy that "well-funded investigators" (defined as $750,000 per year) will not be funded unless program staff make a case otherwise.

Application size: Dr. Beverly Torok-Storb endorsed the idea of reducing the length of an application. She said that, when she was chair of a study section in the 1980s, the section did an experiment in which reviewers read the abstracts and specific aims of every grant. The resulting rankings were in total agreement with the rankings of the reviewers who had done full reviews. Dr. Scarpa said that more people reading a shorter application might be more valuable than only a few people reading a longer application. Shorter applications might also mean that more of them can be discussed at a meeting. Dr. Torok-Storb said the downside of unscored applications is that reviews go back to the applicant without sufficient information for a revision.

Dr. Edward Pugh praised Dr. Scarpa's presentation for its broad view of the problems and possible solutions faced by the review process. He shared a concern that shorter applications might result in people submitting more of them. However, shorter applications might highlight innovation more clearly than longer, detail-laden applications. He noted that shortening the application would be a drastic change, and NIH would have to advance with appropriate deliberation.

Dr. Leinwand suggested that, referring to previous papers, rather than providing large amounts of detail, might result in shorter applications. She then focused on how to provide incentives to serve on study sections, especially when doing so may mean a member's own applications are reviewed elsewhere. She suggested a working group to address this issue.

Dr. Brenner noted that R21s are limited in the number of pages, but not much time is saved because the applications usually come with appendices. Shorter applications are harder on the applicant because it is hard to determine what a reviewer will consider important to include. A shorter application can take as long to write as a longer one. He also said that shorter applications might result in more reviewer subjectivity. Dr. Scarpa said that the NIH Extramural Advisory Working Group recently discussed the disconnect between what is on an application and the criteria used in reviews. For example, innovation is often buried within an application.

Dr. Anne Sassaman asked for more information about asynchronous reviews. Dr. Scarpa said that biophysicists in the pilots particularly liked these reviews; some of the more "traditional" reviewers liked them and some did not. Reviewers have more time in Internet reviews, but there is a trade-off in not having a face-to-face meeting. He stressed the need to create and offer different platforms, including asynchronous reviews, as a way to get the best reviewers possible.

Shortening the Review Cycle

Dr. Eileen Bradley, Chief of the Surgical Sciences, Biomedical Imaging and Bioengineering IRG, reported for the trans-NIH committee that she chairs to develop a pilot in which amended applications can be submitted for the very next cycle. The committee recommended that the pilot begin in a few study sections for new investigators' applications. If successful, it could expand to

new investigators in all study sections and finally to all applicants in all study sections.

Under the current system, applications that come in on February 1 can be resubmitted on November 1. The shortened cycle would allow for resubmission on July 20. Meeting this time frame requires (1) reengineering processes in receipt and referral, (2) posting summary statements ASAP, and (3) enhancing communications with program staff. Applications would have to be easily fixable and then returned to the same study section. Notifications to all those affected as well as the development of an appropriate evaluation instrument will be critical to the success of the pilot. Finally, a shortened cycle requires resources to make it happen. Dr. Bradley continued by reviewing a timeline for a shortened review cycle for Type 1 and Type 2 applications.

Implementation for Receipt and Referral
Within CSR's Division of Receipt and Referral, meeting this timeline would require separate receipt dates for new investigator proposals, with the amended proposals then auto-assigned rather than validated and referred again. Referral officers would have to expedite the processing of these applications to the IRGs and the windows for the receipt of applications submitted under a request for applications (RFA) or program announcement reviewed in an institute (PAR) would need to be relaxed to accommodate the extra stress. The changes would require additional temporary help and some shifting of current staff.

Implementation for Review
For this shortened cycle to work, Dr. Bradley explained, study sections would have to meet between May 15 and June 15. The periods to send applications to reviewers and for reviewers to review them would be compressed, but reviewers would have the applications for at least one month. Summary statement releases of new investigator applications would be expedited, and there would have to be active communication with program staff.

Forty study sections will be in the pilot, and their meetings already were or have been changed to fall within the May 15-June 15 window. Internet Assisted Reviews will be utilized more fully, and reviewers will have to post their critiques before their study sections meet. The edit phase will be limited—critiques will either be edited at the meetings or that same night. The Division of Extramural Activities Support (DEAS), which supports the SRAs, only provides files after the edit phase ends, which means that edited critiques cannot trail in after a meeting.

Dr. Bradley said that a shortened cycle means there will be emphasis on triangulation among the applicant, review, and program staff. Results from the review will go to program staff, who must then talk with applicants about whether to come back for the next submission. Dr. Philip Smith, Co-director of the Office of Obesity Research, National Institute of Diabetes and Digestive and Kidney Diseases, will chair the Triangulation Committee responsible for communication between review and program staff in the pilot.

Notifications
Notifications about the pilot through various means are very important, including the *NIH Guide* and in program notifications. Investigators will receive automatic e-mails via IMPAC II. In addition, a note will be attached to summary statements of eligible new investigators.

Evaluation
Evaluation is key to determining whether the shortened review cycle can work. An Evaluation Committee, chaired by Dr. Bettie Graham, Associate Director of Extramural Research, National Human Genome Research Institute, has identified stakeholders and timelines, and an outside expert will be responsible for the evaluation instrument.

Resources Needed
Dr. Bradley reviewed resources needed for the pilot, which include temporary staff in Receipt and Referral. The IRGs will need to be creative and streamline using IAR. The IRG chiefs involved in the pilot will get together to look at resources and strategies. A forum of SRAs participating in the pilot has also been launched. She underscored that communication with the SRAs is critical.

Discussion
Evaluation: Dr. Pugh requested that PRAC have the opportunity to see the evaluation instrument and suggested several issues the evaluation should address. He asked how to judge the quality of the resubmitted application on the shorter cycle versus on the longer scale. The issue, he stressed, is not just whether the cycle is shorter but whether the process improved as a result. He also suggested an evaluation of costs and praised the triangulation process, which he termed critical for success. Dr. Bradley said that the Extramural Program Management Committee (EPMC) wants to see the instrument and that the committee would also be happy to share it with PRAC. The pilot committee is made up of many EPMC members with evaluation expertise, and they have addressed many evaluation issues.

Dr. Scarpa said the shortened cycle was important because of the fast pace of science and because it could be a "career saver" for some investigators, especially those seeking tenure or who need bridging funds to keep their labs open.

Resubmission guidance: Dr. Faye Calhoun thanked the committee for its hard work. She asked whether applications from all new investigators would be invited back for the next cycle, and whether it should be the study section, rather than program staff, that invites applicants back on the shortened cycle. She also wondered whether the second summary statement might be shorter. Dr. Bradley said that, after much discussion, the committee decided it was not the role of the study section to talk with investigators about quick resubmission, but rather program staff. One reason, she said, was that the same reviewers might not serve on the next study section. As for shortened summary statements, Dr. Bradley said that this idea was suggested but it could not be implemented if different reviewers were involved the second time. Dr. Leinwand urged the committee to reconsider its decision not to have the study section make a recommendation about a quick resubmission. Even if the reviewers were different the next time, a recommendation would give additional data to the program staff.

Applicant reaction: Dr. Torok-Storb praised the effort of the committee but wondered if the proposed process really addressed applicant's problems. Applicants, she said, want money sooner, but a shorter review cycle does not make more money available. Dr. Scarpa said that the

ultimate goal is to identify the best research, so the earlier that a decision can be made, the better. Dr. Torok-Storb replied that she was concerned that applications would "stack up," and that applications coming in for the first time would be disadvantaged. Dr. Bradley replied that Congress, the NIH Director, Dr. Elias Zerhouni, and much of the extramural community want the opportunity for quicker resubmission.

Dr. Ramm expressed concern that triangulation between the program staff and the applicant might create false expectations of funding. Dr. Bradley stressed that quick resubmission would be presented to applicants only as an opportunity if they wanted it. Dr. Scarpa said that shortened cycles, especially if extended both to new and senior investigators, would represent great value to the scientific community. Dr. Ramm agreed, but expressed concern about the workload.

Dr. Mochly-Rosen said that shortening the resubmission cycle might decrease the number of people who apply for more than one grant. More importantly, she said, it will make a tremendous difference to new investigators. After a transition period, she thought the workload would even out. Dr. Ruiz Bravo stressed that investigators will not know whether or not they will be funded, but rather will know their priority score earlier.

Communication: Dr. Craig McClain stressed the shortened cycle would require an education process for new investigators and institutions, in addition to NIH. Dr. Brenner said that it would be important that applicants who choose not to resubmit in the next cycle would not be prejudiced when they do. He said that young investigators must recognize that amending an application requires a lot of effort, beyond just rewriting. Dr. Sassaman said she wanted to emphasize the issue of communication. It will be a challenge for program staff to get enough information from the study sections to provide to investigators. Dr. Martinez said that the shorter cycle would mean major changes for the review process, with the base of the problem the amount of money available to CSR. Dr. Scarpa replied that the review-cycle decision should not be budget-driven. In the pilot, the major work is in referral and will only affect a small percentage of applications.

Dr. Mochly-Rosen observed that the quicker resubmission will be self-selecting. Investigators who have to make major changes or rethink their applications will not choose to do it. Those who only need to make minor revisions will know sooner if they are funded. She said that for a starting assistant professor, a shorter review cycle is a lifesaving event. Dr. Pugh warned that new investigators might be pressured by their chairs to resubmit earlier, and he said that program staff must know not to push early resubmission to people who are not ready to do so.

Dr. Torok-Storb said the process would put a burden on program staff and asked to hear their views. Dr. Ramm said she wondered if program staff would have the time to provide quality feedback to applicants. Dr. Bradley replied that the pilot will provide more information. She noted that program staff would be meeting with the Triangulation Committee, but agreed that the shortened cycle represents doing business in a different way. Dr. Ruiz Bravo said that the shortened review cycle represents one aspect of how NIH is trying to help new investigators.

Update on Electronic Submission

Ms. Megan Columbus, NIH Program Manager for Electronic Receipt of Grant Applications, updated PRAC members on electronic submission, which she said is becoming a reality after many years of planning. The goal remains to be fully electronic for all applications by May 2007.

NIH's strategy is to transition mechanism-by-mechanism. Once a mechanism's date for transition comes up, all applications for it must be submitted electronically via Grants.gov on a Standard Form 424. Those not yet transitioned still require paper submission on a PHS 398 form. Plans for each mechanism are announced in the NIH Guide for Grants and Contracts.

In December 2005, applications for Small Business Innovation Research (SBIR) and Small Business Technology Transfer (STTR) and R13/U13 applications were transitioned to electronic submission. R36s and R15s are due electronically in February 2006.

Ms. Columbus said that the system has worked. Not only has the electronic process operated as planned, but also the applications themselves are clearer and less grainy. In this first attempt, NIH provided flexibility in submission dates to accommodate system errors. She noted that Grants.gov held up well, with no reports of degraded service, even though submissions to NIH set a new Grants.gov daily record on December 1, 2005, when 1,259 applications were submitted. NIH and Grants.gov have developed a strong working relationship. Ms. Columbus noted that NIH is the largest granting agency that Grants.gov will service, and the two entities are helping each other to meet their goals.

The guiding principles of electronic submission are not to hold applicants responsible for system problems but to expect applicants to complete required registration processes prior to submission and to complete the two steps of submission in a timely manner.

More than 300 people at NIH are working on various aspects of electronic submission. An Electronic Application Coordination Group helps maintain the link among them.

Lessons Learned from December
Ms. Columbus reported on the December electronic submissions. The Helpdesk was overwhelmed, so Helpdesk statistics are being analyzed to figure out necessary improvements. One reason was that electronic submission is new for all NIH staff, so just about any question was referred to the Helpdesk. In-house training and discussions will help bring staff up to speed so they can answer many questions themselves without a Helpdesk referral.

Another lesson learned was the volume of changed/corrected applications was very high. Many of the errors resulted from not following the instructions exactly, for example in not registering both with Grants.gov and the NIH Commons. Some system quirks also occurred, which are being resolved. Validation requirements are being reevaluated, recognizing that automated validations may be catching errors that went unnoticed in the paper process. For example, they are looking at turning some of the error messages into warnings, so that applicants can make corrections without completely resending an application. A new Application Guide was released with 15 pages of additional information and guidance, and there are many online tools.

Outreach with the applicant community continues. Users have provided valuable feedback through a dedicated e-mail address (NIHElectronicSubmission@mail.nih.gov) and electronic submission Web site (http://era.nih.gov/ElectronicReceipt/). She asked PRAC members to help spread the word about electronic submissions.

Discussion
Dr. Calhoun praised the hard work that has gone into electronic submissions. She suggested publishing a hard copy to summarize the process for those who prefer a paper format. Ms. Columbus said that there are brochures and other information online, but budget makes it difficult to print sufficient numbers. A staff member from the National Institute of Allergy and Infectious Diseases (NIAID) has offered to synthesize the Application Guide. Dr. Calhoun stressed the need for a manual that simplifies the process.

Dr. Matt Winkler shared comments from the grant writer in his organization about electronic submission. The grant writer found December to be frustrating, but recognized the process would go smoother in April. Ms. Columbus sympathized and agreed that many people were frustrated, especially when technical problems could not be resolved quickly. She said that she received a generally supportive reaction at a recent meeting of the Federal Demonstration Project.

Dr. Pugh asked whether a dummy submission could be made so applicants could check their applications. Ms. Columbus said that a similar tool should be ready by April. She and her associates are also working with larger institutions that do system-to-system development so they can check validations before submitting an application to Grants.gov. Dr. Pugh asked about the scalability of the process and a "Plan B" contingency. Dr. Ruiz Bravo praised Ms. Columbus' coordination and leadership of this effort, and assured PRAC of contingency plans. Dr. Pugh suggested documenting the process so that it can be shared with others.

Review of Biomedical Grants by the U.S. Army

Colonel Janet Harris, Director, Congressionally Directed Medical Research Programs (CDMRP) of the U.S. Army Nurse Corps, was invited to speak about peer review in the Department of Defense's CDMRP. She first provided some background on CDMRP, which was created to find and fund the best research to eradicate cancer, particularly breast cancer. The CDMRP was established in 1992 as a result of the advocacy efforts of the National Breast Cancer Coalition.

Innovation Focus
CDMRP has six unique features: (1) funds are added to the Department of Defense (DoD) budget by Congress; (2) the program responds to targeted guidance from Congress; (3) two-tier formal review of proposals (peer review to determine scientific merit followed by programmatic review which identifies the studies recommended for funding; (4) consumer advocate participation throughout the process; (5) the program vision is adapted yearly to facilitate rapid changes and fill gaps; and (6) there are highly flexible management processes.

Colonel Harris reviewed CDMRP's funding history, pointing out that breast cancer research remains the highest-funded program followed by research into prostate cancer. The hallmark of

CDMRP is to fund highly innovative research. All programs within the CDMRP depend upon yearly congressional appropriations. After the congressional appropriation has been signed into law, each program's Integration Panel – an expert panel of scientist, clinicians, and consumer advocates – meets to deliberate issues and concerns unique to the individual program and establishes a vision and investment strategy for the coming year. Proposals received in response to published program announcements are subjected to a two-tier review derived from the 1993 Institute of Medicine recommendations. The first tier is a scientific peer review of proposals against established criteria for determination of scientific merit. The second tier of the review process is programmatic review, which is conducted by the Integration Panel. Programmatic review is a comparison-based process using published programmatic review criteria. Scientifically sound proposals that most effectively address the unique focus and goals of the program are recommended for funding. COL. Harris noted that the process has been entirely electronic since 2001. Eventually the system will transition to Grants.gov. COL. Harris also noted that electronic submission improves efficiency and saves money for both CDMRP and applicants.

Colonel Harris then described the CDMRP's peer review and scoring processes. For most grants, such as the Idea Award, proposals receive a global score, similar to the scales used at NIH from 1.0 (best) to 5.0 (worst), as well as a global adjectival rating. Each mechanism also has criteria scores, with criteria rated from 1 (worst) to 10 (best). The global and criteria scores were purposely designed to flow in opposite directions because reviewers are to weigh certain criteria more than others. The Concept Award, designed to fund highly innovative new ideas, requires a brief application with a one to three-page proposal body, and reviews are blinded. COL Harris also discussed the review processes used for a few of the CDMRP's more unique award mechanisms, such as the Innovator Award (similar to the NIH Pioneer Award), Era of Hope Scholar Award (for younger investigators); and Clinical Trial Award.

Lessons Learned

COL Harris reviewed the lessons learned by CDMRP during their 14 years of service. Consumers (disease survivors) are an integral part of both the peer and programmatic review processes. Both the scientific and programmatic reviews are organized by program, which facilitates integration of disease specific survivors for each program. Tremendous efficiencies have been gained with cradle to grave electronic management, even with some growing pains and occasional difficulties. Being as flexible as possible and tailoring the peer review panels to the specific mechanism help to address the goals of each mechanism and time and budget constraints. Ongoing challenges include finding the best way to identify innovation and keeping peer reviewers engaged, during all discussions. Maintaining congruency among the narrative, criteria scores, and the global score is also an ongoing challenge. Lack of congruence makes funding decisions by the Integration Panel (programmatic review) more difficult. COL Harris said she is looking for ways to optimize panel size, reviewer workload, and length of panel discussions, as well as ways to ensure that critiques meet the needs of the Programmatic Review Panels and the applicants. She invited PRAC members to the visit the CDMRP Web site (http://cdmrp.army.mil).

Discussion

Innovation focus: Dr. Martinez asked about the size, length of time of funding, and length of applications for the mechanisms, as well as who provides the global scores and makes funding decisions. COL Harris responded that the size of the award, length of the application and length of funding vary widely depending on the award mechanism. For example, the concept awards have brief applications (1-3 page proposal) and the awards are up to $75,000 direct costs for a 1 year performance period. While the application for the Idea Award, has a 10 page proposal body with awards for up to a total of $375,000 in direct costs for a performance period of up to 3 years. The Integration Panel for each program conducts programmatic review and establishes the recommended for funding list. CDMRP does not use a pay line. The Integration Panel reviews all scientifically meritorious studies and selects the proposals that are both scientifically meritorious and best meet the goals of the program. COL Harris noted that this year applications were presented to Programmatic Review Panels in order of the score received on the Innovation criterion. Dr. Martinez also asked about a post-grant process to measure innovation. COL Harris agreed that this measurement was hard to determine, explaining that the Breast Cancer Program funded the initial herceptin research. She said that the program is predicated on giving new innovative ideas an opportunity to move forward, so they can get funding elsewhere if they show promise. The Breast Cancer Research Program evaluated how many times Concept Award investigators were able to use their results as data for another award.

Dr. Martinez also wondered about the difficulty in recruiting reviewers if they need to focus on innovation rather than on the scientific aspects of an application. Col. Harris said that many of the reviewers also conduct reviews for NIH, and they sometimes express frustration during the out-briefs at the end of a CDMRP panel. Dr. Brenner said that he has served on both NIH and DoD panels, and there are many other "cross-overs." He expressed some difficulty in adjusting to the DoD mechanisms. While praising DoD for a well-organized system, he saw a lack of a study-section dynamic as a weakness, in that DoD panels are more ad hoc and meet once a year. COL Harris agreed that a lack of standing panels is a challenge.

Dr. Mochly-Rosen inquired how many grants have eventually resulted in break-throughs and prominence. She stressed the importance of this test particularly as NIH tries to make innovation a bigger factor in funding decisions. COL Harris said she does not have those data. Dr. Mochly-Rosen noted that people perceive innovation very differently, such that David Kaplan has suggested that deviation of scores from the mean actually indicates real innovation. COL Harris noted that rating and funding innovation is an ongoing challenge and currently under study.

Advocacy groups: Dr. Sassaman asked about CDMRP's interaction with advocacy groups and Congress. COL Harris said that CDMRP does not lobby Congress and that advocacy groups are told they must go to Congress if they want funding for a certain program. She reiterated that CDMRP can only fund what Congress specifically states. A consumer representative group makes recommendations on how to better educate consumers before they become involved in the peer review process. In terms of funding via Congressional appropriations, COL Harris said that they have 2 fiscal years to obligate funds. Once the funds are obligated, the funds are dispersed over a maximum of 5 years. Since all CDMRP programs operate on yearly appropriations, the studies are funded for the entire performance period with that year's appropriation.

Dr. Ruiz Bravo asked whether DoD has something similar to the NIH Computer Retrieval of Information on Scientific Projects (CRISP) system that posts abstracts of funded grants. COL Harris said that information on the grants is available on the CDMRP website. In response to a question from Dr. Torok-Storb, COL Harris explained that the Integration Panel decides which mechanisms will be offered each year in the various programs. For example, last year, in the Breast Cancer Program, the Clinical Transitional Research Award focused on chemotherapeutic prevention.

Public Members on NIH Review Committees

Ms. Christina Clark and Dr. Marjorie Mau, members of the NIH Director's Council of Public Representatives (COPR), were invited by Dr. Scarpa to talk about the concept of public members or public representatives on NIH peer review panels. The speakers provided background materials as important context to their presentation and the discussion, which were made available to PRAC and other meeting attendees. These documents included baseline understandings and premises about the qualifications and roles of public members in peer review, as well as a statement of COPR's purpose and goals in bringing the topic to the PRAC.

Dr. Mau began by further clarifying COPR's interest in a dialogue with PRAC. She confirmed that COPR recognizes peer review at CSR and NIH as the gold standard in the field. Within that context, she acknowledged some of the challenges discussed earlier in the day, such as the need to recruit and train the best reviewers; to effectively evaluate a broad range of clinical research; and to increase the system's transparency, accountability, and uniformity. Based on what COPR has learned from its previous collaborative work on appropriate involvement of the public in the NIH research process, and noting the expressed challenges in peer review, COPR was initiating a dialogue with PRAC and seeking a mechanism to continue it.

Study Participant Experts
Dr. Mau noted that terminology is still emerging, but for now the term "study participant expert" (SPE) would describe the proposed participants. The intent would be a carefully defined and structured involvement in reviews of research applications involving human subjects (not basic research); and the purpose of involvement would be to provide value-added perspective of the target population to be recruited for studies, which would ultimately help achieve research goals.

She emphasized that the SPEs would need to be carefully selected, and trained or experienced in addressing issues such as the adequacy of the recruitment plan, relation of the study design and requirements to participant retention, adequacy of human subjects' protection, and representation of minority populations.

Ms. Clark further described that SPEs would not be self-selected, but rather would be chosen by the SRA because of their specific experience and background through a careful screening process. As with any new reviewer, they would receive formal training or individual orientation on the NIH peer review system and procedures, as well as on the key requirements of conflict of interest and confidentiality. They would be governed by the same rules and regulations as all other reviewers. She also emphasized that SPEs would not advocate for or against any particular research area during the review meeting, and they would not represent an advocacy organization or cause. An SRA would determine whether a particular review would benefit from SPE

expertise, and would also select the specific SPE to participate. The Consumer Advocates in Research and Related Activities (CARRA) Program of the National Cancer Institute is one of several models of SPE involvement in peer review at the NIH. CARRA staff provides customized lists of screened, trained SPEs to SRAs, based on the specific needs of their review; then the SRA selects one or more individual(s) as needed.

Ms. Clark noted that COPR became interested in this subject through a series of its previous reports (linked on their web site), and previous discussion on various NIH initiatives including the NIH Roadmap (especially "Re-engineering the Clinical Research Enterprise") and the NIH Public Trust Initiative, as well as hearing about the Trans-NIH Dialogue on Public Members in Peer Review. She concluded by proposing that PRAC consider a pilot of SPE involvement in peer review meetings. She noted that the COPR meeting in April will be an opportunity to continue the dialogue on this concept.

NIH Experience with Consumer Representatives
Dr. Olivia Bartlett, Chief of the NCI Research Programs Review Branch and this year's chair of the NIH Review Policy Committee (RPC), gave a follow-up presentation in her role as RPC chair. She shared results of a survey sent to all 27 IC review chiefs about whether and how they include consumer advocates in their peer review of clinical research.

Of the 16 review offices that responded to the survey, four indicated that they routinely involve consumer advocates in their clinical reviews: the National Institute of Allergy and Infectious Diseases (NIAID), the National Cancer Institute (NCI), the National Institute on Drug Abuse (NIDA), and the National Institute of Mental Health (NIMH). Three of the ICs that responded indicated they occasionally or sporadically include consumer advocates in reviews: the National Institute of Neurological Disorders and Stroke (NINDS), the National Center on Minority Health and Health Disparities (NCMHD), and the National Institute of Child Health and Human Development (NICHD).

Of the respondents indicating they do not use consumers, the reasons varied. Some have a basic research portfolio. Two ICs said that they do not because diverse constituencies in their areas represent competing interests. Other reasons mentioned included: nurses or clinician assistants are recruited to represent the human subjects' perspective; the IC had had a poor experience with involvement of consumer advocates in the past; or the IC did not think advocates would have the adequate expertise.

Dr. Bartlett reported that the ICs which include consumer advocates in peer review refer to them by different terms (e.g., consumer advocates, community reviewers, public members, etc.), but all have the same goal: to provide the perspective of the patient or target population to be recruited for a study. Dr. Bartlett further explained how the CARRA Program works at NCI.

Overall, advocates must have experience beyond their own personal situation and perspectives. SRAs at NCI receive a list of qualified CARRA member names in response to their CARRA request. SRAs at NIAID, NIDA, and NIMH generally identify potential SPEs through attendance at meetings, recommendations, Internet searches, and other means. NCI, NIMH, and NIDA have formal training programs for SPEs who will participate in peer review; SRAs in all

four ICs provide both standard and adapted review instructions to SPEs. In each case, the SPEs focus on the portions of an application that deal with human subjects issues or community outreach. She related examples from her own experience as an SRA in which SPE involvement provided useful value-added information and perspectives to a review.

Discussion

Patients' perspective: Dr. Brenner, who has served on several NCI committees with SPEs, said that the selection and training of SPEs are critical. He explained that, though clinicians are able to see the patient perspective in grant reviews, the perspective of consumer advocates could also be valuable. He suggested that CSR consider the concept. However, he expressed concern that advocates must understand the nuances in scoring patterns. Dr. Bartlett noted that, in both her previous experience at NIAID and her current experience at NCI, SPEs understand the scoring guidelines and their scores are in line with those provided by scientific reviewers.

Dr. McClain agreed that there are situations in which an SPE can play a major role in ensuring that barriers to participation of human subjects are addressed, though he agreed with Dr. Brenner that the physicians planning the research study as well as those on the review panel would understand patients' concerns. He asked how SPEs score grants if they are not involved in the science, to which Dr. Bartlett responded that they listen to the discussion and score the applications in a manner similar to other reviewers on a review panel; for example, unassigned scientific reviewers also generally take their lead from the discussion and the recommendations of the assigned reviewers. She stressed that SPEs focus on human-subject issues rather than, for example, the science of a specific type of cancer. They are also looking carefully at plans for recruitment and retention of patients.

Dr. Ramm said it was an interesting idea, but that she agreed with the limits expressed by Dr. Brenner and Dr. McClain. She said NCRR includes what they call "research subject advocates" on peer review committees to look at clinical trial protocols and address other human subjects issues. She suggested that other ICs might also want to include this type of reviewer.

Scoring issues: Dr. Mochly-Rosen said that she interpreted the fact that SPEs scored the same as scientific reviewers as a bad sign. Input from the community is important, but advocates should weigh in only on the topics on which they are knowledgeable. She also inquired if penalties existed for advocates who did espouse a particular cause. Dr. Bartlett replied that this was rare in her experience, and that, as with any member of a review committee, if someone appears to have a personal agenda, the SRA takes the person aside to address more appropriate behavior during the review meeting. She stressed that this topic is discussed when SPEs are recruited and asked Col. Harris about her experience at DoD. Col. Harris said she could provide a reference for an article that compared the scoring of SPEs and scientists on CDMRP panels. She also noted that in the DoD reviews, SPEs have separate scoring criteria, and primarily score on relevance and provide valued input on identifying potentially far-reaching ideas. Dr. Brenner concurred that his experience with advocates on DoD panels has been positive.

Dr. Calhoun said that the question comes down to who else besides researchers need to comment on the merit and feasibility of a research proposal and termed these people "OTRs" or "other than researchers." She said that, for her Institute, the most valuable OTR would be a treatment

provider with day-to-day responsibility for a large group of patients. Dr. Bartlett noted that some ICs routinely include a nurse, treatment provider, or similar person, and that often these reviewers could be considered SPEs rather than scientific reviewers. Dr. Calhoun asked Dr. Mau and Ms. Clark if that concept was in line with what COPR was proposing. Dr. Mau said that some of the CARRA members attending the training session she observed included providers. The CARRA members in the training session recognized that they would not be able to or be expected to address most of the technical aspects of proposals, and that they could be ask questions of the scientists on the review panel if they needed clarification of technical issues.

Ms. Clark referred to Dr. Zerhouni's observation that the primary impediment to good medical treatment is compliance. Inclusion of those who represent human subject populations could be of value to NIH in this area by identifying during peer review potential barriers to compliance in clinical trials.

Dr. Pugh said he agreed with the recognition throughout NIH that stakeholders need to be represented throughout the process. However, he referred to the core values of peer review: scientific and technical competence, and fairness and objectivity. Though consumer advocates have tremendous value to add, they may not meet the scientific and technical competence principle. Training and SRAs have been proposed as filters, he said, and advocates provide valuable input, but he does not believe that they should vote. Dr. Torok-Storb related her experience with CARRA members on site visits. She said they were valuable for discussing Institutional Review Board (IRB) concerns. In addition, she noted that as a result of their involvement in site visits and other review experiences, consumer advocates are more likely to become advocates for scientific research, which is important for NIH.

Beyond reviews: Dr. Ruiz Bravo agreed with the importance of the patient perspective and thanked COPR for bringing this issue to PRAC. However, she said that the study section structure makes the concept difficult to implement. Rather than insert the patient perspective at the time of the review, she said it might be more valuable earlier in the process when investigators are preparing their applications. She suggested a compendium of best practices that applicants could use at that time. COPR and consumer advocates also have a valuable role in the funding process. She stressed that she did not want to devalue their input during the review process, but by then, they are only critiquing applications, rather than influencing applications before they are even submitted. Dr. Mau said that COPR would discuss the issue at its next meeting in April.

Dr. Scarpa concluded the discussion by thanking the speakers and noting that the concept would be discussed within several extramural committees. He reflected on the previous discussion, saying the main concern seemed to center on scoring by SPEs.

Update on Priority Score and Criterion Scoring

Dr. James Onken, Chief, NIGMS Office of Program Analysis and Evaluation, reported for a working group formed by the Office of Extramural Research (OER) to follow up on PRAC recommendations on scoring. Two main issues are whether additional information from reviewer

scores can assist principal investigators (PIs) or staff, and how decision-making research can benefit peer review.

Measures of Dispersion
The working group looked at possible information contained in reviewers' scores beyond priority and percentile scores that could be better taken advantage of by PIs or staff and, if so, how to present the information.

Issues with reporting score dispersions: The working group examined distributions of individual reviewer scores from a sample of applications reviewed by CSR study sections, to see if there are alternative methods for characterizing the distributions and presenting the information. Given the broad variety of study sections and patterns of scoring, they concluded no numerical measure of dispersion adequately represents score distributions and that graphical displays are a straightforward way to present information about score distributions.

The working group found five categories of problems associated with numerical measures of dispersion: (1) they can fail to capture differences in the underlying distributions; (2) they may not work well with small study sections; (3) they may not be comparable across applications reviewed by different study sections; (4) they can suggest differences that are not otherwise apparent; and (5) they may not be grasped intuititively.

Dr. Onken said the working group discussed how informative measures of dispersion are in the absence of more context about the reasons for the different viewpoints. The group also was concerned about reviewer confidentiality, especially in small study sections, if individual scores are presented in some kind of distribution. Group members questioned whether knowledge about the distribution of individual scores would have detrimental effects on reviewer behavior in unintended ways.

Next steps: The working group proposed three steps:

- Assess the benefits and costs of dispersion reporting, including clarifying the need, usage, and potential benefit of these scores to PIs and staff; exploring any unintended consequences; estimating costs; and determining if there are sufficient benefits to proceed;
- Consider a pilot with a volunteer panel or IC, which would require changes to the information management systems to collect and present scores;
- Consider a "virtual pilot" by using scores from previous reviews and developing mock-ups to present to a group of PIs and staff to see if they find the information useful.

Decision-Making Research in Peer Review
Dr. Jane Steinberg, Director of Extramural Activities, NIMH, said that she would summarize how decision-making research can aid peer review and then brief PRAC members on what the working group learned on this topic since its last presentation to PRAC in September 2005.

Recap: She summarized what decision research is, noting that people turn to certain heuristics or biases to try to simplify complex decisions, even though it may lead to non-optimal outcomes. These biases include accessibility, framing, and anchoring, but they can be minimized or used to

strengthen decision-making. Peer review, which is a set of decisions, has many areas in which decision-making research may apply. Consultants to the Rating of Grant Applications (RGA) subcommittee suggested that breaking down complex decisions improves accuracy. Thus, they recommended that reviewers address each criterion, rather than identifying overall strengths and weaknesses, and this recommendation has been implemented.

One approach might be to score by criterion, either averaging all the criteria scores or weighting them. She acknowledged that many people would not like this system, but it is standard business in contracts review and is what reviewers are implicitly asked to do. In looking for any examples within grant reviews at NIH, she found that one NIGMS panel considered scoring by criteria but ultimately decided not to. She said this reluctance underscores that changes in scoring have as much to do with behavior change as with scientific calibration.

Next Steps: Dr. Steinberg suggested starting with something "small and do-able," such as ensuring that all reviewers get consistent instruction sets. She noticed that a difference in scoring can result when reviewers are asked to "find something important in an application" versus "find something wrong." In addition, it would be useful to collect data on the incremental value of criterion scoring by talking to NIH contractors and program officers who are doing it in their work. Experimental data could include asking two groups to review applications holistically and asking two other groups to undertake criteria-based reviews, and to then see if the correlation between the two holistic scores is higher, the same, or lower than the correlation between the two criteria groups. She underlined that consent and privacy issues, as well as use of reviewers' time, would have to first be carefully considered.

Discussion

Dr. Pugh agreed that incremental steps are the way to move forward in getting more information beyond overall score from the criteria. He added that determining the dispersion of scores could be useful for the study section chair and SRA before a meeting to help decide whether to triage or discuss an application. Dispersion might also be valuable after a meeting when information goes to the ICs for Councils. In these cases, a table with dispersion of scores might be helpful.

Separating out scores: Dr. Leinwand asked whether looking at Program Project Grant (PPG) scores might help determine whether to look at individual criteria scores. The score of an overall PPG does not necessarily match up with the individual scores of the projects, and she asked how often this actually occurs. Dr. Steinberg did not have exact numbers, but noted the policy that allows particularly good projects to be parsed from a PPG. In the case of R01 applications, Dr. Leinwand wondered whether a weighting factor could favor certain criteria over others.

Dr. Ramm said that reviews of P41s with different areas, such as a project or service component and training, have individual scores that reviewers weigh in their minds before deciding on an overall score. Sometimes the mix of the five different scores does not come close to the final score because of different weighting.

Dr. Martin asked PRAC members whether they thought a graph of reviewers' actual priority scores would be useful to applicants. Dr. Pugh thought it might be dangerous because applicants might identify individual reviewers. Dr. Torok-Storb and Dr. Ramm said they thought the

information could result in council letters from applicants. Dr. Sassaman spoke from a council and program staff perspective, saying any additional information is helpful, especially when making funding decisions in a time of tight budgets. She favored looking at options and pilots and offered to discuss being involved with her Institute.

Dr. Pugh said he could think of several visual ways to present the information to a council. Getting information to the decision body is critical, but it is harder to say what to do for the investigator. However, he advised playing it straight, because the system would break down fast if investigators learned the additional scoring information was presented in the meetings. Dr. Ruiz Bravo agreed that these issues have to be worked through and thanked the working group.

The Grant Application Appendix: Who Wants It and What Will We Do with It?

Dr. Suzanne Fisher, Director of the CSR Division of Receipt and Referral, introduced what she termed the fairly practical topic of appendix materials and the opportunity to reconsider this topic as electronic submissions are instituted. A committee looked at the content, format, and delivery of appendix materials to NIH, as well as the provision of these materials to reviewers and storage as part of the official grant file.

She reminded PRAC members about the types of materials currently allowed under PHS 398 instructions for R01 applications, then turned the floor over to Dr. Kirt Vener, Chief of NCI's Special Review and Logistics Branch. Dr. Vener explained that the committee's first step was to determine the value of appendix material to reviewers and program staff. He briefly shared findings of discussion with NCI and CSR study sections on their views about the value of appendix materials, and provided the complete findings in a separate handout. Reviewers generally found appendix materials more valuable than did program staff, but it depends on the mechanism and science. There was little enthusiasm in either group to eliminate appendix materials.

Appendix Materials in Electronic Submissions
Issues to consider in the transition to electronic submissions relate to publications, figures, and other appendix materials. Dr. Fisher suggested that applicants submit publicly available URLs of articles whenever possible, and PDFs only when not possible. She also proposed a maximum number that would vary be mechanism. Smaller images for figures can now be included since they can easily be magnified. Other types of materials are mechanism-specific. Dr. Fisher said that perhaps the protocols now in an appendix of a clinical proposal should be in the main application. How to include special media, such as computer games or sample heart valves, will continue to be discussed with SRAs on a case-by-case basis. Standard instructions on sending PDFs are being developed. Ultimately, applicants may be able to send their appendix materials separately and at a later date.

Appendix materials were sent to small business reviewers on a CD in December, but other options being considered include making them available on a secure Web site. NIH also needs to archive appendix materials as part of the official grant file. For now, some are stored

electronically and some on paper, though eventually most materials will be stored electronically. She concluded by welcoming PRAC feedback.

Discussion
Dr. Calhoun said that it seems the committee has thought through most of the issues, even if they do not yet have all the answers. She asked where collaboration letters go in the submission, and Dr. Fisher said they are part of the application and are submitted as PDFs.

Size: Dr. Torok-Storb said that some items can be eliminated, such as extra images and reprints. Other items are still important, such as protocols and IRB materials. Overall, she said she favors significantly reducing the size and content of appendices. Dr. Winkler said that in this "post-Google world," an appendix becomes a guide to point reviewers to pertinent knowledge. Dr. Brenner stressed the need not to impose limits on protocol documents. Survey tools are also integral to population-based research and must be part of the application. Dr. Scarpa speculated that different needs may be in place for clinical and non-clinical research. He said his concern was "reviewer fatigue," when reviewers have too many additional pages to read. Dr. Ramm said that, while she liked the idea of looking at reducing appendices, it might not be possible with some larger mechanisms.

Ease of review: Dr. Pugh said that the process should be as easy as possible for reviewers. Having publications available without having to find them on the Internet facilitates a review. Dr. Vener said that instructions must be provided, citing instances when applicants have sent encrypted or password-protected files to NCI. He acknowledged the security concerns that some applicants have, especially in SBIR areas. Dr. Berg observed that URLs to articles that require a fee to access could be a problem for some reviewers.

CSR Peer Review Outcomes for Human Subjects Research Applications

In 1994, the "Williams Report" [Williams, G. H., Chair. 1994. "An Analysis of the Review of Patient- Oriented Research (POR) Grant Applications by the DRG, NIH Clinical Research Study Group"] described differences in peer review outcomes between patient oriented and non-patient-oriented research applications. On average 21 percent of non-patient oriented research applications scored in the 20th percentile or better, as compared to 18 percent of patient oriented research applications. More recently, Dr. Ted Kotchen, CSR Special Advisor on Clinical Research, looked at a subset of CSR study sections and reported a similar outcome when using human subjects research identifiers as a surrogate for clinical research. However, Dr. Martin noted that the use of averages can be misleading. He presented examples of how variations in the relative rates at which different subgroups of applicants submit applications, while the peer review outcomes for each subgroup remain the same, can result in different overall averages.

Dr. Martin noted that for an average group of R01 applications reviewed by a CSR study section, one would expect 20 percent of the applications would score at the 20th percentile or better, 40 percent at the 40th or better, and so forth. However, a subgroup might fare better or worse than that norm. Dr. Martin described a new analysis using the NIH human subjects research definitions [excluding applications designated Exemption 4]. Applications using human subjects were designated HS+ in this study, and those not using human subjects [including those with

exemption 4] were designated HS-. Dr. Martin then presented a graph that plotted all scored HS+ and HS- R01 applications reviewed by CSR in October 2004, January 2005, and May 2005. They included Type 1 and Type 2 applications, HS+ and HS-, from new and from established investigators. This graph demonstrated that type 1 HS+ applications from new investigators fare worse than those type 1 HS- applications from new investigators; type 1 HS+ applications from established investigators fare worse than those type 1 HS- applications from established investigators; and, type 2 HS+ applications fare worse than type 2 HS- applications.

Initial Applications (A0)
A further analysis was conducted looking at the effect of initial and revised applications. An analysis of Type 1 A0 (initial) applications submitted by new investigators over 12 council rounds from October 2000 through May 2004 showed that HS- applications fared better than HS+ applications at the 20th percentile, with considerable round-to-round variation. Experienced investigators fared slightly better, but again Type 1 HS+ applications did not fare as well as HS- applications at the 20th percentile. Type 2 applications fared better than Type 1 applications, again with considerable dispersion from round to round. Summarizing his findings of initial applications, Dr. Martin observed that there are differences between HS+ and HS- within each type of application and across types of A0 applications.

Revised Applications (A1 and A2)
Moving on to A1 (revised) applications, the difference between HS+ and HS- applications at the 20th percentile substantially narrowed, with a dramatic improvement in outcomes for HS+ applications. For A1 applications the differences in outcomes for experienced versus new investigators did remain, as did the advantage of a Type 2 versus a Type 1 application. Looking at A2 applications, while fewer in number, showed no difference between HS+ and HS- applications. The advantage of a Type 2 application was gone at that point.

Questions Raised
Dr. Martin said the findings raise additional questions, for example: Why do HS+ applications improve so much on revision? And, do HS+ PIs submit applications at the same rate as HS- PIs?

The data indicate that type 2 HS+ applications are submitted at a lower rate than Type 2 HS- applications, which he termed disturbing. To confirm this finding, Dr. Martin looked at Type 1 applications funded between 1996 and 1999, and broke them down by new and established PIs and by HS+ and HS-. Both new and established HS+ investigators' awarded a type 1 grant are less likely to submit a type 2 application than those awarded grants that are HS-. He queried why, as well as how this difference affects the reported difference in global review outcomes. Before turning the presentation over to Dr. Kotchen, Dr. Martin acknowledged the staff who helped with the analysis.

Differences in Clinical Scores
Dr. Kotchen addressed some of the issues of why HS+ applications seem to have less favorable outcomes than HS- applications. He began by listing what are <u>not</u> reasons: the density of HS+

applications considered by a study section, costs associated with clinical research, composition of study sections, or whether an application is considered by a CSR or an IC panel.

Instead, potential contributing factors may include whether the application raises safety or privacy concerns during the review. Another factor may be the training of clinical investigators, as data show that individuals who have participated in a clinical research-training program have fared better in getting subsequent funding. Review criteria may make a difference, and he noted that the criteria were revised about a year ago to be more supportive of clinical research. Finally, he questioned whether clinical research is more difficult, with more hurdles for investigators.

Discussion
Dr. Brenner characterized the data as extraordinary and important, and shared his own views of reasons behind the findings. He said issues of training are involved, as well as increasing financial pressures on clinicians. He urged Dr. Martin and Dr. Kotchen to publish their analysis.

Dr. McClain expressed concern about the rising average age of a first-time R01 applicant who is an M.D. Dr. Martin observed that new HS+ applicants come back with a Type 1 revised application at a higher rate than HS- applicants, but at a lower rate for Type 2 applications. Dr. McClain suggested that financial pressures may lead M.D.s to leave research for private practice.

Dr. Sassaman wondered if the results would be the same if research conducted by M.D.s involving human subjects were parsed out from research that did not involve M.D.s, and that might not be clinical research. She agreed about the importance of publishing the information to continue to inform the community. Dr. Kotchen said that some of the information has been published.

Dr. Pugh said that the "take-home message" he gained is a hypothesis that might explain why HS+ and HS- applications have fared differently. In addition to learning that HS+ applications tend to improve greatly after revision, he pointed to two other effects discussed: senior investigators do better on average, and perhaps the nature of a clinical application is that it typically has a single grant life-cycle. Dr. Martin said that the fact that clinical investigators may need to start with a new proposal more often than other investigators could be stressful. Dr. Kotchen also speculated that perhaps a clinical study takes longer to conduct, and thus clinical investigators do not have sufficient information to compete for a Type 2 renewal.

Dr. Martinez related that he has heard that treatment grants do better than prevention grants, and it might be interesting to compare these two categories, as well as outcomes at the behavioral versus non-behavioral ICs. Dr. Torok-Storb asked whether there was a way to find out what happens to the investigators who do not come back with Type 2 applications. Dr. Leinwand suggested that perhaps some of the research gets picked up drug companies.

Dr. Ruiz Bravo said knowledge management for disease coding will make more information available. Drs. Brenner and Pugh underlined the need to disseminate the findings presented.

Restructuring Seven Study Sections in Three CSR IRGs

Dr. Don Schneider, Director of the CSR Division of Molecular and Cellular Mechanisms, reported to PRAC about restructuring some study sections. He explained that restructuring responds to various situations. The study sections in question are in three IRGs: Bioengineering Sciences and Technologies, where the issue is small business applications; Cell Biology, where the issue is glycobiology applications; and Molecular Cellular, and Developmental Neurosciences, where one study section is too small and one is too large.

Bioengineering Sciences and Technologies IRG
This IRG was created in January 2004, with the intent to review R01 and small business applications together. Six hybrid study sections were created, but they have become quite large. To handle the load, it is proposed to move small business applications from three of the study sections into four separate panels: Delivery Systems and Nanotechnologies, Materials Science and Environmental Monitoring, Devices and Detection Systems, and Assays and Methods Development. The other three study sections would continue to consider R01s and small business applications together at this time.

Cell Biology IRG
This IRG was reorganized in January 2005. Glycobiology went into the Intracellular Interactions (ICI) and Cell Structure and Function (CSF) study sections, with the involvement of a glycobiologist on the design team and endorsed by glycobiologists. Though ICI is fine as glycobiology's primary home, the secondary home of CSF has not worked well. An ad hoc group at a meeting of the Society for Glycobiology proposed the Membrane Biology and Protein Processing study section as glycobiology's secondary home, endorsed by others in the field.

Molecular, Cellular, and Developmental Neurosciences IRG
The plan for this IRG centers on two of its seven study sections: Neuro-Degeneration and Biology of Glia (NDBG), and Biophysics of Synapses, Channels, and Transporters (BSCT). Over time, NDBG has grown to 150 applications per round, while BSCT has decreased, with some rounds fewer than 40. A working group consisting of representatives of the two study sections and NIH staff realized that the applications could not just be divided between the two sections. They proposed splitting out mitochondrial neuron death applications from NDBG and renaming the study section Neuronal Degenerative Disorders and Gilial Biology. A SEP would be chartered on Neuronal Oxidative Metabolism and Death. The working group anticipates that each panel would review about 75 applications per cycle.

The group observed that the BSCT study section is strong but extremely competitive, which may contribute to decreasing numbers of applications. They proposed a solution that moves both applications and reviewers in and out to change the dynamics, involving three other study sections and renaming BSCT to Biophysics of Neuronal Systems (BPNS). They anticipate that this study section would start with about 50 applications per cycle, growing to 60 to 80 applications per cycle. This realignment would also make more efficient use of reviewers.

Discussion
Dr. Ramm applauded the recognition that knowledge management does not replace dealing with the sociodynamics and culture of study sections. Monitoring study sections is critical to their success. Dr. Pugh concurred, and also praised the effort to involve the glycobiology community.

Dr. Ruiz Bravo asked whether published transparent criteria indicate when it is time to address an issue within a study section or IRG. Dr. Schneider said that these cases arose from attending study sections and talking to reviewers, chairs, and program staff. Dr. Pugh suggested beginning the presentation with the goals of the PSBR process, which would put the specific changes in context. Dr. Ruiz Bravo praised the effort and said that she just did not want CSR to be peppered by requests since there are hundreds of study sections. Dr. Martin added that looking at size is one of the first criteria as a trigger to change. The IRGs were created through the PBSR process, and what they are doing now is tweaking within the IRGs.

Dr. Mochly-Rosen pointed to advantages in reviewing academic and small business bioengineering applications together. She expressed disappointment that they had to be separated out, to which Dr. Schneider agreed and said that the SRAs really wanted to make it work, too.

PRAC members agreed that the proposed changes in the study sections of the three IRGs discussed should go forward.

CSR Review of Member Conflict Applications

Dr. Samuel Edwards, SRA of the Cellular and Molecular Immunology Study Section, represented a working group composed of Drs. Anshumali Chaudhari, Valerie Durrant, Marc Rigas and himself, tasked with looking at the process for dealing with member SEPs. He expressed gratitude to the Query Viewing and Reporting team and the IRG chiefs on behalf of the working group.

Dr. Edwards reviewed the law pertaining to member SEPs: when a peer review group meets regularly, the applications of its members must be reviewed by another qualified review group to ensure a competent, objective review. Current law does not define a "different review group," though previously it stated such a group could have no more than 50 percent of members of the original study section.

CSR policy is that member applications can be reviewed in another appropriate standing study section if one exists or, if not, in a member SEP. The SEP has to be less than 50 percent current or recent members of the conflicted member's standing study section, and the SEP chair cannot be a member of that study section. If 30 to 50 percent of the SEP are current or recently retired members, then the application is percentiled against the parent study section. Normally, if there are fewer than 30 percent, the application is percentiled against the CSR All database.

During the 2005 council rounds, most (816) member applications were reviewed in other study sections, but a large number (691) went to member SEPs. Over the past 6 years, the percentages have been about the same, even as total numbers increased. A survey of IRG chiefs showed that member applications are reviewed in a variety of ways, and 95 percent of the IRGs use multiple formats driven by appropriate expertise. Most IRGs review some or all member applications in groups (sister study sections or SEPs) that do not contain any current or recent members of the conflicted study section. One concern is that reviewers seem to carry the mentality of the study

section from which they came in terms of percentile scoring, so they are sometimes not on the same playing field when the newly constituted group comes together.

Most member SEPs are now held about a month after the study section. Most take place by phone to consider between one and five applications. A new directive has recommended a minimum of three applications for a member SEP. The composition of these SEPs reflects standing study sections in terms of seniority, so most are full professors. As an aside, however, Dr. Edwards observed that this analysis showed an upward trend in the participation of associate and assistant professors in both standing study sections and SEPs, and a downward trend in the participation of full professors.

The meat of the analysis, said Dr. Edwards, was to compare how R01 member applications reviewed in CSR fared in member SEPs, regular study sections, and CSR All reviews. He noted that analyses had been conducted in 1991 and 1999, and it is good to revisit this issue periodically. They looked for the percentages of applications in the top 10 percent, the top 20 percent, and the bottom 70 percent, including streamlined applications. They also looked at any differences between Type 1 and Type 2 applications, as well as those that did and did not use human subjects. Dr. Edwards presented a number of graphs that addressed these categories

Generally, Type 1 applications without human subjects reviewed in a member SEP do better than the other categories. There was no significant difference in scores if a Type 2 application was reviewed in a study section or a member SEP, but both fared better than those reviewed against CSR All. Fewer applications reviewed in member SEPs are streamlined compared to those reviewed in standing study section meetings. The question also arose if members are at an advantage as a regular member of a study section. The group looked at the year before members joined a study section compared to when they became members, and found they received comparable scores before and during service.

In summary, member conflict applications represent a significant segment of applications reviewed by CSR, because the many members across the study sections are highly productive. At the same time, they provide a service to NIH. The goal is fair review for member applications given that by regulation they cannot be reviewed in their parent study section, which may be the most appropriate venue for their application. There is no systematic bias in the review of their applications, particularly in terms of the use of member SEPs. They do tend to do well, but this is likely because they are excellent scientists who can write good applications. The current system provides appropriate means for reviewing member applications. Flexibility and the role of the IRG chief are integral making the process work.

They suggest that CSR develop a consistent policy about whether current members of a study section can review member applications. They also recommend that CSR reconsider the policy to percentile against CSR All, but rather score against the parent study section. They have concerns about the confidentiality of reviewers in small member SEPs, especially for assistant professors reviewing full professors' applications. They are also concerned about designer reviews and the SRA workload. They are encouraged by the recent changes of scheduling member SEPs closer to when study sections meet and of avoiding SEPs that consider only one or two applications.

Discussion
Dr. Mochly-Rosen thanked the working group for this analysis and stressed the importance of disseminating the findings. She also noted that she counsels assistant professors against serving on small panels because of the issue of being easily identified. Dr. Edwards noted, however, that often assistant professors are the leaders in their field, particularly as technology moves rapidly.

Dr. Ramm concurred that keeping the number of applications in a SEP up to three to five is important. Dr. Edwards said that the balance is sometimes putting some applications into a larger review disadvantages them if they cannot be reviewed appropriately. Dr. Torok-Storb said many of her colleagues feel that the law is based on an incorrect assumption that there is less objectivity within a study section than in member SEPs.

Dr. Brenner commented on a "Catch-22": if there are different reviewers in a SEP than the parent study section, the percentiling falls apart because the member's application is percentiled against CSR All. However, if reviewers from the parent study section serve on the SEP and the grant fails, there could be interpersonal issues that affect how the study section functions. Given these choices, Dr. Brenner said that he would prefer recruiting non-member reviewers for a SEP to keep the culture of the study section intact. Dr. Edwards noted that there has not been a difference between how applications have scored in member SEPs and study sections. The SRA gives the SEP guidance on ballpark numbers of what an outstanding application might score.

Dr. Ruiz Bravo said that the law is designed to address the perception of conflict, even if no actual conflicts occur. Dr. Pugh said that CSR has steered between the need to protect study section integrity and the willingness of people to serve on study sections. Ultimately, he said, the system passes the test of "do no harm." It is running in a way that is working.

Conflict of Interest/Review of Clinical Applications in ICs

Staff from three ICs—NHLBI, NIMH, and NIAID--discussed how they deal with member conflicts in clinical research applications.

NHLBI
Dr. Valerie Prenger, Chief of the NHLBI Review Branch, explained that her IC reviews multi-site clinical trials and CSR reviews single-site clinical trials. Applicants meet with program staff and a review representative before they submit a grant and the review representative informs the applicant how their application will be reviewed so they can avoid common pitfalls.

She described how the standing study section, called the Clinical Trials Review Committee (CLTR), operates. It is made up of 16 experts augmented by ad hoc members. All standing members not in conflict read and can weigh in on every application. About 12 studies per round are reviewed. The typical study involves 2,000 or more patients, at six to nine sites, but may involve as many as 150 sites. Routine conflicts are handled under standard NIH policy. Applications by standing members are reviewed in SEPs, organized and conducted by a different SRA. Current CLTR members are not allowed to participate in a SEP. Past CLTR members who did not serve with the current member are recruited and typically make up about 50 percent of the member SEP. The meeting is held via telephone. Anecdotally, there is no advantage or

disadvantage to a SEP compared to CLTR review. An attempt is made to limit SEP discussion so that an application is considered for the same length of time as in an in-person CLTR meeting.

NIMH

Dr. David Sommers, SRA in the NIMH Review Branch, described the current review structure. There are two standing committees: the Interventions Research Review Committee (ITV) and the Services Research Review Committee. SEPs are held in each round for approximately 10 to 12 applications, including member conflicts, as well to handle RFAs and the overall workload. As of July 1, 2006, a restructuring will result in five smaller standing committees.

Conflicts of interest have not been a large problem because NIMH trials tend to be smaller, although some conflicts can occur with multi-site trials and multiple PIs, or with affiliations with drug companies. The new structure should provide a home for many member conflicts and decrease the need for SEPs. Applications from permanent members are assigned to another standing study section or a SEP. If temporary reviewers have significant roles within an application, they are not invited to that particular meeting; if they have only a minor role, they leave the room during the discussion.

The Conflict of Interest SEPs do not necessarily have parent committee representation, either present or previous. Applications are percentiled against a joint NIMH database. Best practice dictates that the committee's SRA not handle the SEP. Anecdotally, there does not seem to be any difference for an application considered in a SEP or a parent section, although there is a lower rate of streamlining in the SEPs.

NIAID

Dr. Hortencia Hornbeak, Associate Director for Scientific Review and Policy, NIAID, focused on dealing with conflicts of interest in review of two NIAID initiatives developed when the HIV/AIDS Clinical Research Enterprise was re-engineered. Balancing appropriate expertise and managing conflicts of interest has been a challenge, because of the collaborative nature of the clinical networks, product competition, trial oversight support, and the institutions and personnel involved. Most of the research areas and collaborations involved other countries, which adds more complexity and the need for in-country expertise. NIAID found it could not stay within the guidelines for chartered committees and get the appropriate expertise, and so has gone to SEPs.

Currently, there are applications from 58 countries for the RFA "Units for HIV/AIDS Clinical Trial Networks." In total, they involve more than 9,000 personnel and 650 institutions so the conflict of interest challenges are large. NIAID worked closely with the OER to develop review strategies. The applications were divided into 14 SEPs with an electronic system to facilitate the tracking and determination of the conflicts of interest. This database was crucial in tracking potential conflicts.

She presented some examples of the waivers NIAID received. For example, non-key clinical trial unit (CTU) and leadership personnel may review non-affiliated CTU applications. As another example, reviewers may be used who have published with key personnel in an application within the last 2 years (reduced from 3 years).

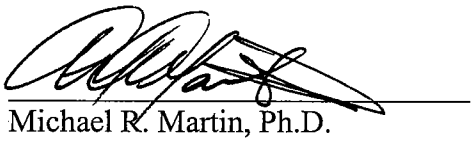
Dr. Hornbeak acknowledged the contributions of OER. She stressed that any relaxations of conflicts of interest be communicated to the review committee at the orientation of the meeting and included in the permanent record. Dr. Ruiz Bravo clarified that these are formal peer review waivers, and no laws were changed.

Discussion
Dr. Sassaman said she found it interesting to learn how different ICs handle conflicts of interest and useful for PRAC to see how ICs are handling different issues. She noted that presentations like these make clear that peer review is an NIH, not just a CSR, responsibility.

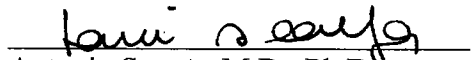
Concluding Remarks

Dr. Martin thanked the participants, and reminded them that the next PRAC meeting is May 22, 2006, 8:30 a.m. PRAC adjourned the meeting at 5:11 p.m.

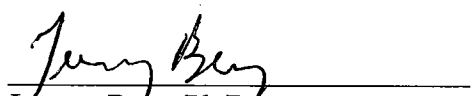
We do hereby certify that, to the best of our knowledge, the foregoing minutes of the January 2006 meeting of PRAC are accurate and complete. The minutes will be considered at the May 2006 meeting of the Advisory Committee, and any corrections or comments will be made at that meeting.



Michael R. Martin, Ph.D.
Executive Secretary
Center for Scientific Review Advisory Committee



Antonio Scarpa, M.D., Ph.D.
Co-Chair
Peer Review Advisory Committee



Jeremy Berg, Ph.D.
Co-Chair
Peer Review Advisory Committee