The role of the study section in the NIH peer review system

BERNICE SACKS LIPKIN

*Computer and Biomathematical Sciences Study Section, Division of Research Grants
National Institutes of Health, Bethesda, Maryland 20014*

This paper describes the NIH grants review system, which is designed to evaluate a grant proposal for both technical merit and programmatic interest. In particular, it discusses the initial review for scientific and technical merit, using as example the Computer Study Section, which reviews applications in computer science, physiological modeling, and statistics. Criteria for evaluation include: knowledge of the state of the field, relevance to scientific innovation, commitment by the investigator, strength of association between theoretician and practitioner, methodology, human factors, and experience.

THE REVIEW PROCEDURE

The complex of the grants review mechanism at NIH is a multiprocessor, distributed system, whose input is new or renewal applications, some 11,500 entries per year, and whose output is, using fiscal 1975 as reference year, some 5,500 paid grants at a cost of 318 million dollars.

Your application arriving at NIH is first processed by the Referral Branch. The required budgetary and book-keeping details are checked and then health science administrators assign the application to both a study section and an institute. It is the Referral Branch that decides which of the approximately 50 study sections is the appropriate primary review group to assess the technical merit of the proposal. Applications superficially similar in that they propose to model aspects of the cognitive process, say, may receive different study section assignments, depending on the emphasis and direction of the proposed project. If the emphasis is on the development of a theoretical model with heavy use of computers, the application may well come to the Computer and Biomathematical Sciences Study Section, of which I serve as Executive Secretary. A somewhat similar proposal, which also uses a computer for data analysis but which is primarily concerned with the development of experimental procedures for data collection, would more likely go to Experimental Psychology.

As you may know, NIH is mission oriented. It is by law directed to fund applications with a bearing on biomedical and medical research, with the aim of furthering health care. And this is an instruction we take very seriously. NIH is not intended, as is the National Science Foundation, for example, to administer pure science grants. In practice, as you know, few proposals are pure anything, either pure science or pure application. Hence, it is not uncommon for an investigator to submit very similar proposals to both agencies. He usually has the

grace, however, to note somewhere in the NIH version how his project will contribute, say, to the curing of cancer. It is up to the Study Section to determine whether he may not, indeed, be right.

The Referral Branch also assigns the application to an institute that will fund the grant, if it is approved. Assignment is made on the basis of the direction of the proposed project and where it would best fit a particular program of one of the various institutes that comprise the National Institutes of Health. Each institute has an Advisory Council, which reviews the recommendations of the initial review groups. Members of the Council may, if they choose, "flag" an application, opening it up to discussion; Council may decide that the application should be re-reviewed by the same or another study section or that the recommended budget should be reduced. What it cannot do is upgrade an application disapproved by Study Section on scientific grounds. Applications not singled out are voted for "en bloc," whereby (and this applies to the large majority of applications) the initial recommendation is accepted. It is then up to Institute Staff to negotiate with the applicants and to commit the Institute to fund the approved grants, based on priority score and programmatic considerations, until funds for that fiscal year are exhausted.

Thus, the actual merit review is two-tiered: The Study Section reviews a grant for technical merit and the Institute Advisory Council examines it from the viewpoint of programmatic considerations. This hierarchical system has a turnaround time from receipt by the Referral Branch till Council meeting of some 7 to 8 months. The time taken to actually commit funds, because of budgetary uncertainties, may be considerably longer. As an instance, of the grants received by the Computer Study Section during 1974, some 61 out of 94 reviewed, or, in dollar value, some 11 million out of the 42 million requested, have been approved for fund-

ing and 29 have been paid. Of the 1975 grants, only 16 out of the 86 reviewed have to date been funded. (This is not surprising, considering that we have just finished the September Study Section review, so the applications have not yet gone to Council.)

MERIT REVIEW BY STUDY SECTION

The rest of my talk will concentrate on the initial level of review, using the Computer and Biomathematical Sciences Study Section as a representative example.

Types of Applications

The Computer Study Section reviews three types of applications. First are those in computer science. The particular consequence to the Computer Study Section of being mission oriented is that we receive few pure computer science proposals. Vanguard concepts in artificial intelligence, in process control, in multitask, multiprocessor systems are indeed being funded, but it is because they are judged likely to advance biomedical research and thus, ultimately, health care.

There was a time that NIH actively encouraged the installation of that mysterious unknown, the computer, in medical schools to upgrade data collection and expedite data analysis in biomedical research. That time is past. We are now in an era where there is a wealth of modules for building systems. We have modular sensors more or less easy to interface with microprogrammed microprocessors, minicomputers, maximinis, minimaxis, numbercrunchers, and networks of larger machines, each with its own set of satellites. Thus, some of computer science has become an art form, where the components are mostly available but the real issue is how best to integrate them for a particular set of tasks. The emphasis is on reliability, exportability, speed, accuracy, cost effectiveness. Elegance is appreciated, but not as much as is a demonstrably clean system that will do the job when confronted with real sticky dirty data.

To give you an idea of some of the projects that are coming in for review, we are beginning to see some good process control systems in chemistry. These show concern with quality, and use microprocessors or miniprocessors as preprocessors of the raw data prior to transmission to a larger machine. There are also new and interesting graphics configurations to display three-dimensional information, such as the central nervous system reconstructions Professor Levinthal showed us this morning. The EMI tomographic scanner for noninvasive viewing of brain tissue is already beginning to have a profound effect on both two-dimensional picture processing and three-dimensional reconstruction. It is an important new tool in medical diagnosis. Its reconstructions have life and death consequences, and the reconstructed pictures are definitely not of superior photographic quality. These are, as you might surmise, all conditions that would be expected to encourage new

work in reconstruction and enhancement algorithms.

How to handle large data bases and registries is a problem that is not solved in any general sense. It is a problem that will not go away, given the projected demands of national health insurance and the more immediate need for accurate data for demographic analyses. The contradictory requirements of privacy and freedom of information have only increased the magnitude of the number of factors that must be considered. Similarly, medical diagnosis continues to be an important goal. Procedures and concepts from the field of artificial intelligence are just beginning to be applied to model the diagnostic and decision making processes involved in patient management.

The second major area reviewed by the Study Section is modeling, i.e., the modeling of physiological systems, etc. Here there is the perennial problem of delicately balancing between so much simplicity that the model is too weak to be predictive and so much realism that the model is intractable. In some cases, Study Section recognizes that even when dozens to hundreds of simultaneous differential equations are solved using highly sophisticated programming, the results are, if not trite, then just as predictable without the use of the computer. Indeed, when we request opinions from clinicians or researchers, quite often they comment that the investigator has done nothing with the computer that could not have been done without the machine. And they are right, just as those who bet on the Pony Express to outrun the early steam engine were right in the short haul. By analogy, Study Section recognizes that in several areas really innovative modeling is going to require a great deal of development: The problem here is to identify those groups who show sustained progress toward a very difficult goal.

Statistics is the third large category. Many of the proposals deal with the exigencies of clinical trials in cancer chemotherapy—incomplete data, trials prematurely terminated, the inability to balance treatments in a classical fashion. Others have a direct but not immediate application to biomedical requirements. The heuristic used appears to be: the more innovative the approach, the more the distance from immediate application that is tolerated. Contrariwise, if the method has already been shown to be feasible, or if what is proposed is really just manipulating another parameter in a technique already known in some detail, then Study Section will demand vigorous proof that the proposed project has important applicability to a particular problem and that the investigator and/or his collaborators show deep understanding of the clinical or biomedical problem for which the proposed method purports to be useful.

Study Section Membership

The members of the Study Section, i.e., the peers in the peer review system, are chosen most obviously for their technical expertise. We cover a considerable

territory, and study section members are as a group knowledgeable in statistics, in computer science, in modeling particular systems, and in the content areas of patient management, radiology, clinical pathology, neurophysiology, and so forth. And since any field can be further subdivided, we supplement this expertise with reviews from a spectrum of experts in the sub-sub-sub categories involved.

Technical expertise is a prerequisite, but in itself it is not sufficient. Members will perforce be making ethical judgments based on the informed opinion in the field, that is, within the context of values of the medical and biomedical disciplines involved.

The third characteristic of Study Section membership is group judgment. Indeed, if there is anything that characterizes the NIH peer review system, it is the on-line, real-time, interactive, concurrent hand-shaking, finger-waving interchanges that go on during the meetings in which proposals are evaluated. It is a time for fine tuning, for putting the weights on the components that make up a proposal. The computer science is evaluated. The content area is evaluated. The purpose. These are balanced. Is the science terrific but the problem silly? Is the problem important, but implementation finessed? It is a time for plain speaking, for logical thinking, for sober and balanced judgment. Ultimately it is group decision, a distributed evaluation which has encompassed much agonizing and deliberation. To paraphrase the Gestaltists, the summary is greater than, less than, a welding together of and yet different from the sum of its parts. The Summary Statement becomes the primary document of the assessment of the grant. All that has come before, the outside opinions, the preliminary estimates of the members, comments made during discussion, are now, if not irrelevant, then certainly incomplete and in many cases nonpredictive of the group determination. And always there is the hope that the taxpayers' money, their money, is being spent as wisely as mortals are capable of judging what is wise.

Criteria of Technical Merit

With this as background, allow me to explain what we mean by "technically good."

Obviously, an application ought to be well documented and show an awareness of the literature. It is folly to propose to reinvent the wheel; it is the ultimate in folly to propose it when the original inventor is a Study Section member. Incidentally, you might find it a useful addition to your library to write NIH for a copy of the *NIH Public Advisory Groups*, which has rosters of the memberships of all NIH Study Sections, Advisory Councils, and Committees.

To repeat what was said earlier: The proposed research ideally ought to be both highly innovative and very applicable to a particular area or to a set of related problems. To be neither will clearly not increase the probability of being funded.

It is important that the investigator be committed to the project. It must be his project, something he wants to do. To appear to have a shopping basket full of tasks that the investigator is willing to work on, depending on what the funding agency is interested in supporting, is not perceived as a sign of intense commitment.

In applications that straddle both technique and content area and most of ours do—the strength of the collaboration between engineer and clinician, between theory and practice, becomes an important element in the evaluation of the grant. It is recognized that, too often in the past, elegant engineering approaches flourished because the biologist was not in the driver's seat but was used only as a source of information. Unfortunately, most of the subtle and hard-to-articulate biomedical information thus was not communicated and the results were often disastrous.

This is not to say that the application can ignore implementation. I do not recall an application that did not present itself as having an important goal. I recall many that were silent on the ways and means of implementing the goal, yet this is perhaps the most important part of the application. You ought to discuss the constraints to the problem. Where do you expect to encounter difficulties. What strategies and alternative approaches do you have available—your contingency plans. It is pointless to belabor some new-to-you idea that is not really new and then not show how your way of implementing this idea differs from previous approaches.

Human factors are important. In general, a device intended for paraprofessional personnel requires more, not less, elaboration of safeguards and concern with possible misuses or unintended uses than does a device operated by its originator, who is on call for instant debugging and reprogramming. If it is a computer facility, the amount of hand-holding, the degree to which the facility is committed to making it convenient for the researcher, is of concern. Documentation of equipment and scrutability of operator manuals are important because they increase the exportability and widespread use of the system.

The investigator's track record is important to the extent that if he has done difficult and similar projects, Study Section is more confident in his ability to perform in the future. Contrariwise, if an engineer with no biological expertise suddenly decides to branch out into modeling the heart, he would be well advised to persuade a sound clinician to act as consultant, or better still, active collaborator.

Study Section is aware of institutional guidelines for different program areas. For example, facility grants from the Division of Research Resources are expected to have several components: core research, service, collaboration, and training. Each facility has its own personality, with varying emphasis on the components. We take these individual differences into account, while taking note that requirements are at least covered. In

cases of a new program, we may request guidance from Institute Staff who attend the Study Section deliberations.

Ethical considerations are of concern. It might appear that work primarily involved with some aspect of computers could ignore present concerns with high ethical standards. It has been the position of the Study Section for some time that experimental procedures that involve some modification in standard practices of patient management in order to obtain data or to test a program come under the human subjects category and ought to be approved by the Human Subjects committee at the investigator's institution. We are, however, beginning to wonder, more recently, if the pendulum has not swung too far over, so that it is becoming exceedingly difficult, if not impossible, to do studies, say, on attitudinal systems that often require the subject be ignorant of the conditions of the experiment.

Having considered all these factors, Study Section

then takes a long hard look at the budget. A sensible budget is one which is lean but not malnourished. The investigator who underestimates the costs of a proposed project will give the impression of inexperience and inadequate understanding of what it takes to get the job done. Contrariwise, Study Section is accustomed to revising budgets downward in line with a strong consciousness of cost-effectiveness in research.

The result of the assessment of a grant is a priority score and a Summary Statement, commonly known as a pink sheet because of the color of the paper on which it is printed. The pink sheet describes the project briefly and summarizes what the Study Section considered the strengths and weaknesses of the project. This balance is reflected in the priority score, which will be somewhere between 1 and 5, if the application is approved. The application, its priority score, and its pink sheet are now sent to the appropriate institute for the next step in the review procedure.