Blood Cells, Molecules, and Diseases (2001) **27**(6) Nov/Dec: 1020–1027
doi:10.1006/bcmd.2001.0476, available online at http://www.idealibrary.com on IDEAL®

The Productivity and Impact of The Leukemia & Lymphoma Society Scholar Program: The Apparent Positive Effect of Peer Review

Submitted 12/05/01
(Communicated by M. A. Lichtman, M.D., 12/05/01)

Marshall A. Lichtman[1,2] and David Oakes[2]

ABSTRACT: A study was conducted to compare the "productivity" of a cohort of research grant applicants selected by peer review to be scholars of The Leukemia Society of America (now The Leukemia & Lymphoma Society) with a matched cohort of applicants not so selected during the period 1981 to 1990. One hundred and twenty-four scholars and 124 nonfunded applicants were studied. Two bibliometric variables and their derivatives were examined from the Institute of Scientific Information database: the number of papers published and the number of citations to those papers. Published papers were measured through December 31, 1999, and citation counts to these papers through December 31, 2000. Scholars published 10,301 papers through the period of observation and nonfunded applicants published 6442 papers. Scholars' papers were cited 419,798 times, whereas nonfunded applicants' papers were cited 245,586 times. The mean citations per paper were 52 for scholars and 38 for nonfunded applicants. The papers published per scholar, citations per scholar, and citations per paper per scholar were significantly greater than the corresponding measures for nonfunded applicants ($P < 0.0001$ in each case). Scholar's papers were cited 30% more often, whereas nonfunded applicants were cited 10% more frequently, than a comparison group of scientists publishing in the same journal in the same year. High-impact papers, e.g., papers that were cited more than 200 times, were nearly three times as frequent among scholars (494 papers) as among nonfunded applicants (173 papers). This difference was highly significant. The good (better than baseline) performance of nonfunded applicants may be a reflection of self-selection among the applicant pool for this competitive award; the more productive performance of the scholars is probably the result of the selection decisions made during the peer-review process.

Key Words: peer review; research productivity; citation impact; career development.

INTRODUCTION

Organizations that sponsor biomedical research have few quantitative methods for analyzing the results of their investment (1). In the case of voluntary health agencies, lay board members, many of whom are familiar with business models, sometimes ask for appraisals of "productivity" resulting from the research investment. The issues in arriving at a measure of productivity for this purpose are numerous and complex. The result of an investment often takes years or decades to reach fruition. Several research sponsors such as government, industry, and private philanthropies may fund the research either simultaneously or sequentially. The total number of published papers resulting from a sponsored research program can be enormous and determining their impact is difficult. Achievements like patents and licenses of new drugs or products are relatively infrequent and may be impossible to link precisely, retrospectively, to some of the basic concepts from which they were spawned. The ferreting out of the relative contribution of one or another paper

Correspondence and reprint requests to: M. A. Lichtman, University of Rochester Medical Center, 601 Elmwood Avenue, Box 610, Rochester, NY 14642. Fax: (716)271-1876. E-mail: mal@urmc.rochester.edu.
[1] The Leukemia & Lymphoma Society, White Plains, New York 10605.
[2] University of Rochester Medical Center, Rochester, New York 14642.

1079-9796/01 $35.00

Blood Cells, Molecules, and Diseases (2001) **27**(6) Nov/Dec: 1020–1027
doi:10.1006/bcmd.2001.0476, available online at http://www.idealibrary.com on IDEAL®

among many that by accretion may lead to a discovery is very challenging and time consuming.

In an effort to examine the results of the grant funding program of The Leukemia & Lymphoma Society, we identified the published papers of a cohort of individuals who were selected to receive a scholar award and compared their research paper output and the impact of these papers as judged by their citations to those of a matched group of applicants who were not selected for a scholar award. The Leukemia & Lymphoma Society is a voluntary health agency that has funded research on the diagnosis and treatment of leukemia, lymphoma, and myeloma for over 45 years.

The scholar program is only one of several Society-sponsored research programs. Using the scholar program had several advantages in this study. First, these applicants must have reached early research career independence at the time of review. Thus, most applicants are very likely to have had continued research activities whether they received a scholar award or not. The selection of a scholar is largely related to the potential of the recipient and a prediction of his or her future productivity. Fellow and special fellow awards, which are also a part of the Career Development Program, result in large measure from the standing of the senior scientist–mentor in whose laboratory the fellow works and the short-term productivity is a shared outcome. Also, the scholar has a record of early achievement during research fellowship and the first years of independent work, which makes judgements more tangible than in the case of selection of fellows.

To use quantitative variables that could be recovered from an electronic database, we chose two measures available from publication and citation databases of the Institute for Scientific Information (ISI): number of papers published and the number of citations to those papers (2). We used these data and their derivatives to compare the scientific publications of the scholars to the nonfunded applicants.

METHODS

Study Population

A total of 124 scholars selected consecutively by the Grant Review Subcommittee of the Society's Medical and Scientific Committee from 1981 to 1990 was identified in the Society's grantee database. A comparison group of an equal number of individuals in each year was selected at random from those applicants who were not awarded a scholar grant. The nonfunded applicants in 1 year (1987) could not be recovered from the archives of the Society, so that number of nonfunded applicants was added from the prior (1986) and from the succeeding (1988) year to equalize the number of study subjects and their potential years of productivity in each group. A very small number of nonfunded applicants were replaced by substitute choices because they were awarded a scholar grant at a later date during the period of the study. The year 1981 was selected as the starting date because that is the beginning year of the analytical in-house version of ISI's citation database, which was used for this study. The 10-year period chosen left a minimum of 9 years of observation of papers published after the last set of scholars was chosen in 1990.

Identification of Published Papers

The surname and the first initial or the first and middle initial of the 248 subjects was searched in the ISI database of published papers from January 1, 1981, to December 31, 1999. The ISI database contains several varieties of article types including abstracts. This analysis was performed using articles, reviews, notes, and papers published in proceedings issues of journals. Abstracts were excluded from analysis. Notes are short papers but have not been segregated from articles in the ISI database since 1996. The papers were retrieved and examined in chronological sequence and false matches (papers by other authors of the same name) could be eliminated. The total number of relevant papers authored or coauthored by the 248 subjects over the 19-year period was 16,743.

Citation of Published Papers

The number of citations in the final list of papers was obtained for each paper. The citations were then summarized for individuals in the study

and compared between the two groups: scholars and nonfunded applicants. A measure of "compere group" citation frequency was made also. The latter estimate of relative performance (a baseline) was made by measuring the average citation frequency of papers of a similar type (e.g., article, note, review, or proceeding), in the same journal, and in the same year as those published by scholars or nonfunded applicants. ISI refers to this measure as the "expected citation frequency."

Statistical Methods

Data files were created for each scholar and nonfunded applicant giving the total number of papers and the total number of citations for each member of each group. The mean number of papers, the mean number of citations per paper, and the cumulative frequency distributions of papers per study subject as well as citations per paper per study subject were generated. Where applicable, means, standard deviations, standard errors, and medians were calculated. Log transformation of data was used to reduce skewness sufficiently to use a parametric statistical test, the two-sample *t* test. The Wilcoxon test was also used to compare the distributions of relevant variables between scholars and nonfunded applicants. One nonfunded applicant was omitted from the calculations because no published papers could be found for that individual.

RESULTS

Table 1 indicates that the scholars as a group published 3859 more papers during the period of study, 60% more than did nonfunded applicants. This difference was present for each category of paper: articles, short articles or notes, reviews, and articles in proceedings (Table 2). Scholars and nonfunded applicants had 96 and 93% of their papers cited, respectively. Since the number of papers published by scholars was greater, it was little surprise that they were cited as a group more than twice as often as the nonfunded applicants (Table 3). However, the mean number of papers published per scholar, the mean number of citations per scholar, and the mean number of citations per paper per scholar were each significantly greater than those of nonfunded applicants (Table 4). In this analysis, the total number of papers and citations were greater than shown in Table 1 because the analysis required that a paper coauthored by more than one scholar or more than one nonfunded applicant be fully credited to each individual in the analysis of papers or citations per subject. Four-point-six (4.6) percent of papers and 4.8% of citations had scholars as coauthors and 0.9% of papers and 1.0% of citations had nonfunded applicants as coauthors. This greater collaboration among scholars may be the result of the presence of several scholars at certain large cancer centers.

TABLE 1

Aggregate Publication Record of Scholars and Nonfunded Applicants (1981–1990)

	Scholars (N = 124)	Nonfunded applicants (N = 124)
Number of publications[a]	10,301	6,442

[a] Articles, notes, proceedings, reviews.

TABLE 2

Distribution of Reports Analyzed

	Scholars		Nonfunded applicants	
	No.	%	No.	%
Articles	8,726	85	5467	85
Notes	572	5	326	5
Proceedings	518	5	363	6
Reviews	485	5	286	4
Total	10,301	100	6442	100

TABLE 3

Aggregate Citations to Papers Published by Scholars and Nonfunded Applicants (1981–1990)

	Scholars (N = 124)	Nonfunded applicants[a] (N = 123)
Number of citations	536,283	245,586
Percentage of papers cited	95.9	93.5

[a] One nonfunded applicant had no published papers.

Blood Cells, Molecules, and Diseases (2001) **27**(6) Nov/Dec: 1020–1027
doi:10.1006/bcmd.2001.0476, available online at http://www.idealibrary.com on IDEAL®

TABLE 4

Comparison of Published Papers and Citations of Papers: Scholars and Nonfunded Applicants

	Scholars (N = 124)	Nonfunded applicants[a] (N = 123)	t test[c] P value
Number of papers/subject	87 ± 69 (64)	53 ± 59 (38)	<0.0001
Number of citations/subject	4534 ± 4584 (2902)	2017 ± 2809 (1015)	<0.0001
Mean number of citations/paper/subject	48 ± 25 (44)	34 ± 23 (29)	<0.0001
Number of papers with >200 citations/subject[b]	4.2 ± 6.1	1.4 ± 3.5	<0.0001

Note. Data are means ± standard deviation. Median value is in parentheses.

[a] One nonfunded applicant had no published papers during the period of study.

[b] In the aggregate, scholars published 494 (4.8% of all papers) and nonfunded applicants published 173 papers (2.7% of all papers) that were cited more than 200 times.

[c] t tests were applied to log-transformed data to reduce the effect of skewness.

The frequency distribution of papers per study subject and citations per paper per study subject are shown in Figs. 1–3. The citation frequency per paper was greater for scholars than nonfunded applicants at all citation rates. Notably, papers that were cited more than 200 times were nearly

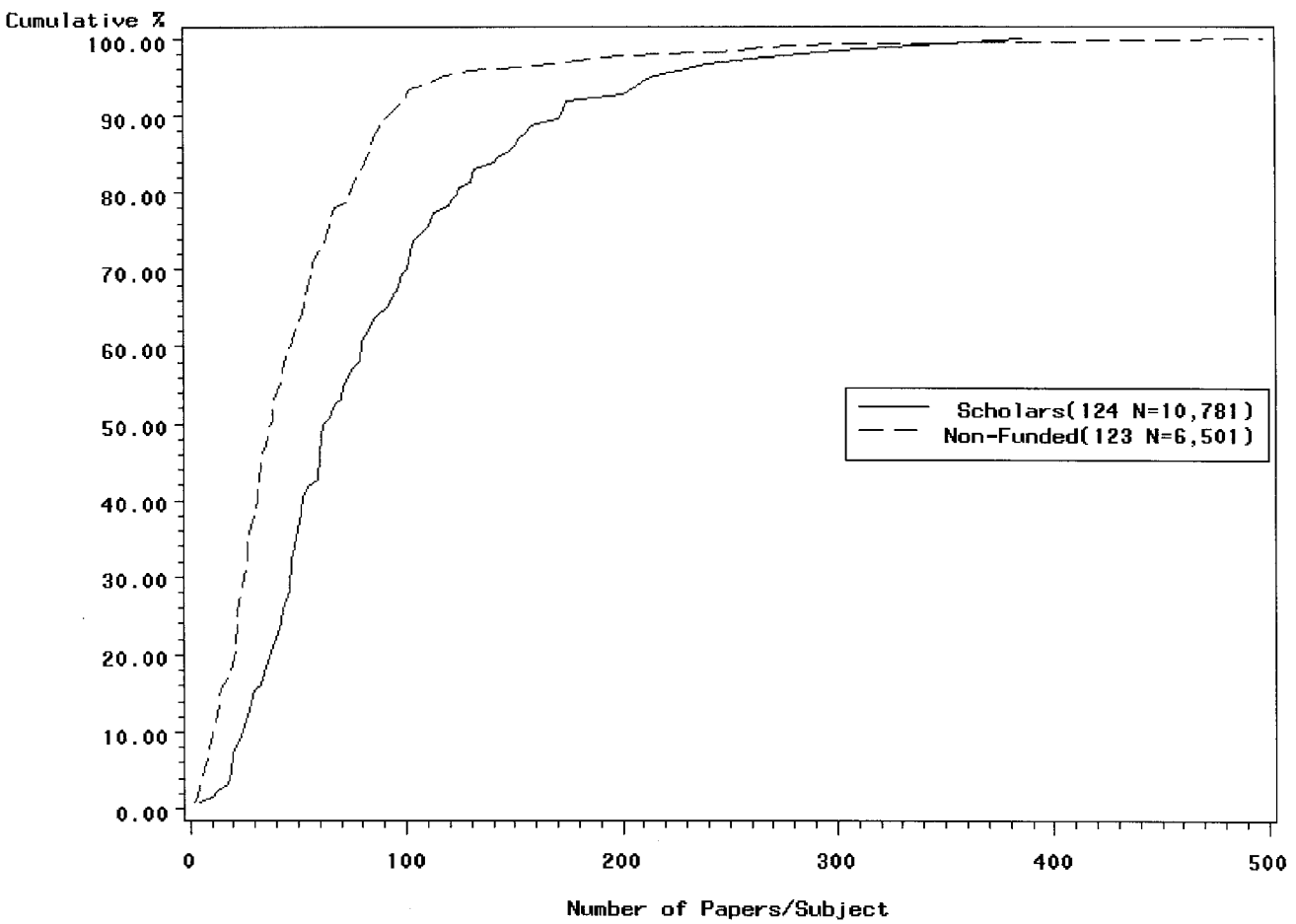

FIG. 1. The cumulative frequency distributions of published papers per study subject for scholars and nonfunded applicants. The number of papers (N) is 4.6 and 1.0% greater for scholars and nonfunded applicants than the total number of papers shown in Table 1. More than one scholar or nonfunded applicant coauthored some papers. In this analysis of productivity per subject, jointly authored papers were credited to each study subject involved. The two curves are significantly different ($P < 0.0001$) by the Wilcoxon test.

doi:10.1006/bcmd.2001.0476, available online at http://www.idealibrary.com on IDEAL®

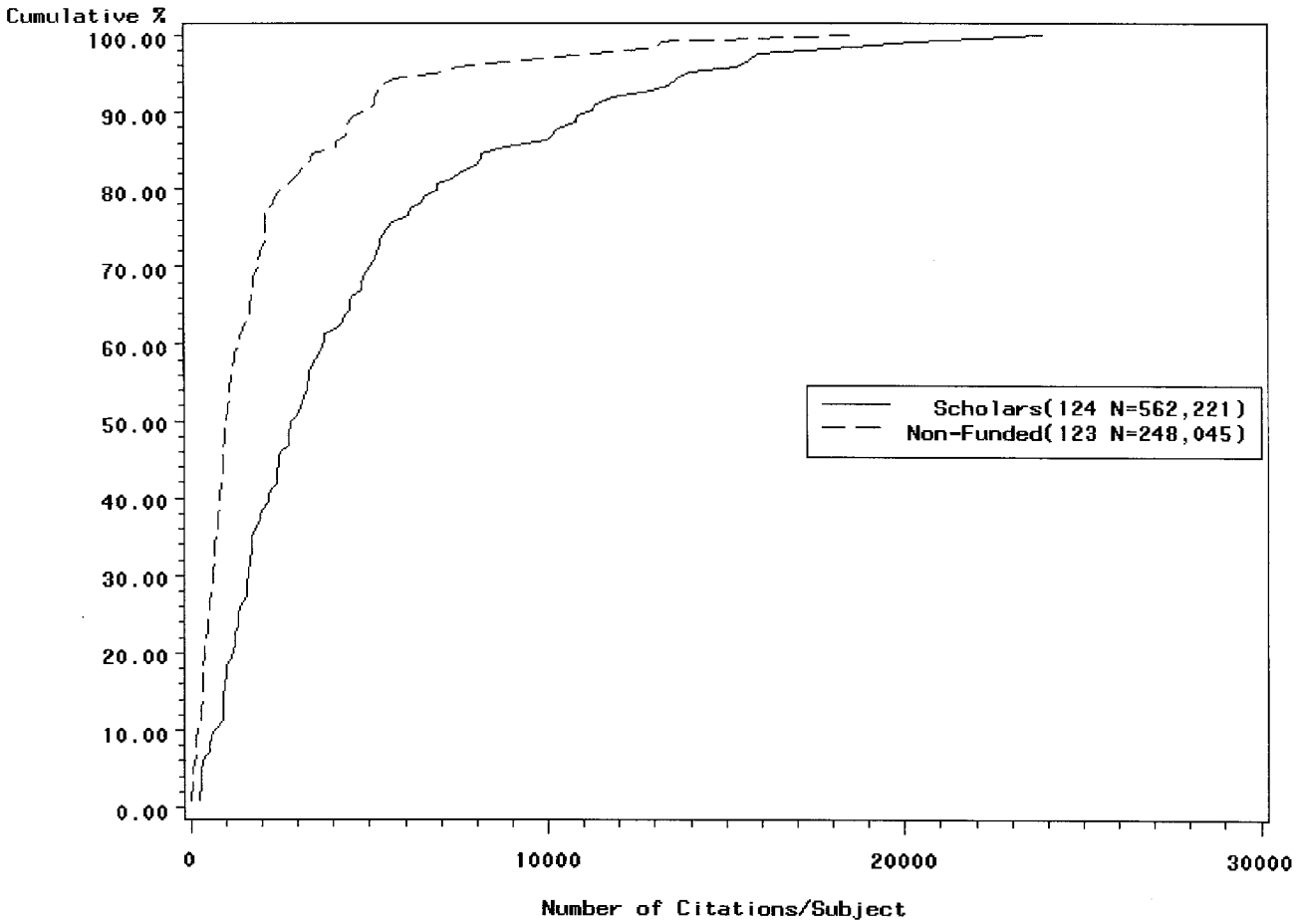

FIG. 2. The cumulative frequency distribution of citations per study subject for scholars and nonfunded applicants. The number of total citations (*N*) is 4.8 and 1.0% greater for scholars and nonfunded applicants than shown in Table 3. More than one scholar or nonfunded applicant coauthored some papers. In this analysis the citations to the jointly authored papers were credited to each author in the study group. The two curves are significantly different ($P < 0.0001$) by the Wilcoxon test.

three times as frequent among scholars as a group (494) than nonfunded applicants as a group (173).

The scholars had a markedly increased number of citations to their papers than a compere group of scientists publishing papers in the same journal at the same time. Scholars were cited 30% more frequently than papers by a compere group of authors. The papers of nonfunded applicants were cited 10% more frequently than papers by a compere group of authors (Table 5).

The cost to the Society of providing a stipend to the 124 scholars, each for 5 years, totaled $8,550,000 over the 14 years of support (stipend ranged from $125,000 to $200,000 per 5 years of award from 1981 to 1994). The periodic increase in stipend, although not precisely linked to the biomedical price index, represents a partial inflation adjustment over this period. Using these expenses, the direct cost to the Society was $830 per scholar publication. In addition, one can estimate the Society's ancillary cost of maintaining a sponsored research program. The cost of administration, fund raising, the grant processing and review process, staff and office space is about 25% of total expenses bringing the cost of a scholar's published paper to about $1100 each. It should be noted that the papers include those published after the Society's support ended but in terms of the Society's mission such subsequent accomplishments count. The total research cost of each paper is unknown since it would include institutional costs and costs expended by other sponsors of the research, notably the Federal government. It can be estimated that this total, even restricted to the same

Blood Cells, Molecules, and Diseases (2001) **27**(6) Nov/Dec: 1020–1027
doi:10.1006/bcmd.2001.0476, available online at http://www.idealibrary.com on IDEAL®

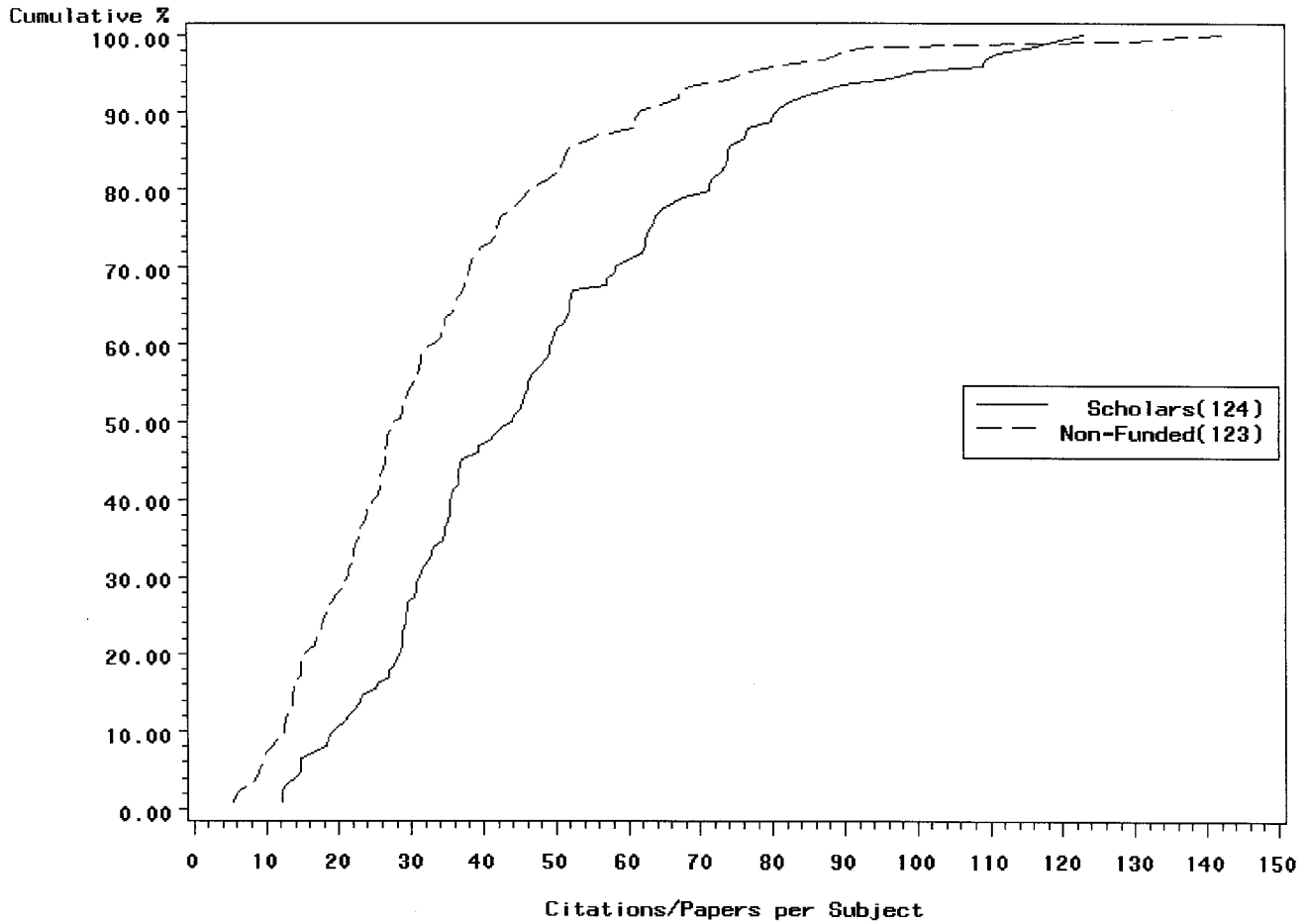

FIG. 3. The cumulative frequency distributions of citations per paper per study subject. The two curves are significantly different ($P < 0.0001$) by the Wilcoxon test.

5-year period of Society's sponsorship, would be between 10- and 20-fold the Society's investment.

DISCUSSION

Although most interested parties agree that the basis of the effort to prevent or cure disease should be an investment in and conduct of relevant scientific research, the process by which limited funds are invested is subject to question. In occasional instances, where an opportunity of singular focus exists, a voluntary health organization may use a czar to drive fund raising and investment for one major treatment objective. This was

TABLE 5

Relationship of the Citations of a Compere Group to the Citations of Scholars and Nonfunded Applicants

Scholars' citations	Compere's citations	Scholars' citations minus comperes' citations	Ratio of scholars' to comperes' citations
536,283	419,798	116,585	1.3
Nonfunded applicants' citations	**Compere's citations**	**Nonfunded applicants' citations minus comperes' citations**	**Ratio of nonfunded applicants' to comperes' citations**
245,586	222,991	23,595	1.1

Note. Citations of the compere group represent the averages of all papers matched for article type, journal, and year of publication to each paper published by the scholars or of the nonfunded applicants.

the case when Basil O'Connor led the National Foundation for Infantile Paralysis (March of Dimes) and provided support of Jonas Salk's work to develop and field-test a polio vaccine. With the success of such an initiative, the justification of that investment is indisputable in retrospect. The singular path to preventing an infection by a small and closely related family of viruses that displayed stable antigenic features made the risk of putting all one's eggs in one basket acceptable. In the case of The Leukemia & Lymphoma Society, the diseases of concern result from somatic mutations that involve dozens and probably hundreds of protooncogenes and result in cancers with widely different biochemical and antigenic targets for therapy. With dozens and perhaps hundreds of neoplastic genotypes and phenotypes, a singular attack is not rational.

In such a setting, it is widely accepted that the best way to determine the investment of a health agency's funds among competing requests is to have a group of independent experts in the field make a consensus decision on the portfolio to be supported, i.e., the peer-review process. Virtually every reputable public and private funding agency uses expert scientific panels to assign priority to competing requests for funding. These two principles—investment in biomedical research and use of a peer-review model to make discriminating decisions regarding the research to be supported—are the bedrock of the effort by health agencies in the United States to ameliorate the effects of disease.

The peer-review process is imperfect because perfection depends on forecasting the outcome of complex projects, determining the most direct route to a clinically applicable innovation, and foreseeing an unexpected discovery. The sardonic descriptions by Nobel Prize laureates of the rejection of their initial paper describing the seminal ideas that ultimately led to the Prize is testimony to the limitations of peer review, at least for journal article reviews. The peer-review process may be generally less inclined to fund ideas that are outside current paradigms or at the edge of the capability of current technologies. This last reservation suggests that the review process is influenced by the radiologist's aphorism: "we see what we know." Compensating for these limitations is the enormous breadth and diversity of the worldwide network of biomedical science, the customary wholesome motives of peer review, the face-to-face discussions that help resolve potential disagreement, the influence of more farsighted peer-review group members who insist on "taking a chance," the willingness of investigators to explore improbable pathways in any case, and the continuing introduction to the mix of new young investigators and reviewers less encumbered with preconceptions.

The program's productivity, as judged by published paper and the impact of those papers as judged by the frequency of their citation among the 124 scientists selected as scholars between 1981 and 1990, was significantly greater than that of a group of 124 nonfunded applicants. There are several inferences that one can derive from these differences. The first is that the peer-review process resulted in a selective enrichment of productive scientists. The applicant pool probably had an important element of self-selection. Persons with strong early career research records are more likely to apply for the scholar award and thus applicants not selected would be expected to have continued research productivity. This is evident from the good publication record and citation frequency of the nonfunded applicants and their somewhat better than expected citation frequency as measured against a compere group. In addition, a small proportion of scholar applicants does not receive awards because they are too advanced in their career. These nonfunded applicants would also increase the productivity of the nonfunded comparison group.

Although there are no unambiguous quantitative measures of research quality in the biomedical sciences, citation frequency has some power. Publication in a peer-reviewed journal is a minimal criterion as an index of quality and this criterion varies by the quality of journal as judged by the standards set by and the competition for publication in that journal.

An alternative interpretation of the findings is that the scholar award provided resources that resulted in or substantially contributed to the research productivity of selected applicants. We

Blood Cells, Molecules, and Diseases (2001) **27**(6) Nov/Dec: 1020–1027
doi:10.1006/bcmd.2001.0476, available online at http://www.idealibrary.com on IDEAL®

have no way of assessing that possibility. The productivity of nonfunded applicants suggests that they garnered continued research support during the period of study although the proportion with sustained careers may have been smaller. Moreover, the scholar award is for salary support (often partial) and requires other research grant funding to pay for additional research costs.

Several treatises have explored the relationship of number of published papers and their citation record to research accomplishment (3–6). Publication rate does predict scientific impact although that correlation explains perhaps about 35 to 50% of the relationship (4). Highly honored scientists are more likely to have higher publication and citation rates than others are. However, some scientists have low publication rates and high citation rates. In judging individuals, the possible combinations of high and low publication rates and high and low citation rates are relevant. Scientists have been categorized into four groups: prolific (higher publication and citation rate), silent (lower publication and citation rate), mass producers (higher publication rate and lower citation rate), and perfectionists (lower publication rate and higher citation rate) (3).

In this study, we were less interested in individual accomplishment than group results. This aim provides more power to the publication and citation rate for the purposes of this analysis. Publication in peer-reviewed journals especially those that are highly cited is one measure of scientific quality. Publication rates and citation rates are strongly correlated (4). Higher citation rates are associated with highly honored scientists such as Nobel Prize Laureates and members of the National Academy of Sciences (5). The reasons to cite a paper are several but the most common reason is thought to be either the cited paper's usefulness to either the methods or the reasoning of the author doing the citing or that it extends the work on that subject.

We conclude that it is probable that selection through the peer-review process resulted in a more productive cohort of scientists funded by the Society than would have been achieved by random selection. By linking its scholar award to the requirement to have additional funding for research costs, the Society contributes to an enormous amount of published results of high impact at a very low investment, a necessity for a voluntary health agency dependent on annual donations to fund its programs.

ACKNOWLEDGMENTS

The authors acknowledge Joanne M. Janciuras for her assistance in statistical analysis and Susan M. Daley for her assistance in the preparation of the manuscript. David A. Pendlebury at the Institute for Scientific Information provided invaluable assistance in the identification of relevant papers and citations and provided them in the form of a database for statistical analysis. He also reviewed and made suggestions in regard to the manuscript.

REFERENCES

1. Luukkonen-Gronow, T. (1987) Scientific research evaluation: A review of methods and various contexts of their application. *R D Manage.* **17,** 207–221.
2. Garfield, E. (1979) Citation Indexing: Its Theory and Applications in Science, Technology, and Humanities. Wiley, New York.
3. Cole, J. R., and Cole, S. (1973) Social Stratification in Science. Univ. Chicago Press, Chicago.
4. Simonton, D. K. (1988) Scientific Genius: A Psychology of Science. Cambridge Univ. Press, New York.
5. Feist, G. J. (1997) Quantity, quality, and depth of research as influences on scientific eminence: Is quantity most important? *Creativity Res. J.* **10,** 325–335.
6. Armstrong, P. W., Caverson, M. M., Adams, L., *et al.* (1997) Evaluation of the Heart and Stroke Foundation of Canada Research Scholarship Program: Research productivity and impact. *Can. J. Cardiol.* **13,** 507–516.