Impact of National Heart, Lung, and Blood Institute–Supported Cardiovascular Epidemiology Research, 1998 to 2012

Richard R. Fabsitz, PhD; George J. Papanicolaou, PhD; Phylliss Sholinsky, MSPH; Sean A. Coady, MA; Cashell E. Jaquish, PhD; Cheryl R. Nelson, MSPH; Jean L. Olson, MD; Mona A. Puggal, MPH; Kevin L. Purkiser; Pothur R. Srinivas, PhD; Gina S. Wei, MD; Michael Wolz, MA; Paul D. Sorlie, PhD

In a recent article, Alberts et al.¹ warned that the US-based biomedical science enterprise is flawed in its assumption that the enterprise will constantly expand, and it cannot expect a persistently expanding National Institutes of Health (NIH) biomedical research budget in the future. In fact, as noted by multiple observers, the NIH budget has been declining in constant dollars since 2003, and the pay lines for grants are at historic lows.¹-⁴ Such realizations have led to calls for a reexamination of the policies and programs of the NIH and its individual institutes and centers.¹,² Some have questioned the appropriateness of the allocation of research budgets between discovery science and translational science.¹,²,⁵ Others have argued for better methods to make funding decisions, suggesting that peer review is too conservative and limits innovation,¹,²,⁴,⁶,⁷ or simply does not demonstrate the ability to prioritize research proposals on their potential to yield high impact.⁸ Evidence that scientists spend too much time writing and rewriting grant applications, stay too long in training programs, and achieve their first tenure track position or first NIH grant in their late thirties and early forties, respectively, suggest there is a mismatch in the supply and demand for scientists that must be addressed.¹,⁵,⁷

Editorial see p 1949

Efforts at the NIH to conduct evaluation and self-examination have already begun. Institutes and centers within the NIH have taken multiple approaches. The National Cancer Institute conducted a workshop to make recommendations on how the National Cancer Institute can transform itself for the 21st century to address criticisms of excess expense, repudiated findings, small incremental gains in knowledge, inability to innovate at reasonable cost, and failure to identify research questions with the greatest merit.¹⁰ Staff at the National, Heart, Lung, and Blood Institute (NHLBI) have proposed the use of results-based accountability for strategic decision making to improve the performance of the Institute and the scientists and projects it supports.¹¹-¹³ Published evaluations and perspectives to date have focused on funding mechanisms,⁸ research designs,¹²,¹⁴ research domains,¹⁵,¹⁶ clinical trials,¹⁷ and research portfolio diversity.¹⁸

In the midst of this evaluation effort to move the NHLBI research portfolio toward “better, faster, cheaper, and more responsive to current needs,”¹³ particular attention has been paid to high-cost clinical studies, ie, clinical trials and epidemiological research.¹,²,¹³ Economic returns of large research studies are being evaluated to determine their value.¹⁹ Clinical trials have come under scrutiny for failing to publish results in a timely fashion or, in many cases, not at all.¹⁷ Barriers have been identified that keep results from achieving the impact expected by the public.¹⁵,¹⁷ Epidemiology studies have been criticized for being expensive and outliving their usefulness¹⁵; epidemiology has also had to respond to critiques regarding redundancy, false positives, relevance, efficiency, and incremental value.⁷,¹²,¹³,²⁰

This report offers an impact evaluation of the cardiovascular epidemiology research supported by the NHLBI over a 15-year period (1998–2012). It attempts to address 3 questions: (1) Over this period, what has been the contribution of NHLBI support to high-impact cardiovascular epidemiology research? (2) What are the trends in the contribution of NHLBI support to high-impact cardiovascular epidemiology research over multiple domains to which epidemiology is expected to contribute, and are there differences in these trends by research domain? (3) Does the contribution of NHLBI support extend to the articles appearing in the most-represented journals publishing high-impact cardiovascular epidemiology research?

Methods

Unlike many of the research areas of the NIH, epidemiology is methods based rather than disease or organ systems based. As such, the discipline often crosses multiple diseases and many health topics. Ascertaining of what is a product of epidemiology research is challenging because reports from epidemiology studies do not always use the word epidemiology in titles, abstracts, or keywords associated with the publications. As a result, a search of epidemiology will not always include all the articles resulting from this discipline.
To address this challenge, the authors developed a list of 10 research domains to which cardiovascular epidemiology is likely to contribute. For each domain, a search of the biomedical literature between 1998 and 2012 was performed using SCOPUS (http://www.elsevier.com/online-tools/scopus) during the week of January 24 to 30, 2014. The search terms used for each of the 10 research domains selected to represent cardiovascular epidemiology research are presented in Table 1. The search within each domain included the search terms specific to that domain plus the additional terms presented at the bottom of the table that were common to all domains. This approach provided a degree of consistency across domains. The last term AND NOT randomiz* was added to eliminate randomized clinical trials from the results. The domain New Biology represents a particular point in scientific progress. As these individual fields within New Biology (defined by search terms in Table 1) become integrated into current knowledge, this domain will either have to be reassessed for validity or noted as being temporally defined when the impact of future funding is determined. Average citations per year were calculated based on the year of publication and the total citations identified through the end of 2013. The results of each search were sorted by average number of citations per year.

In each domain, the top 100 publications based on average citations per year were selected and evaluated for authorship, funding support, and acknowledgments to determine whether the publications involved any support by the NHLBI.

NHLBI support for a publication was based on satisfying at least one of the criteria presented in Table 2. Determination of NHLBI support was most easily identified for those publications citing funding source with the grant number beginning with HL. Publications acknowledging NHLBI or National Heart, Lung, and Blood Institute were also straightforward. Those supported by contract funding were more difficult to identify because of a shift in contract number coding system during this time period (shift from a 7-character code to a 17-character code in n=2004; http://www.hhs.gov/regulations/hhsar/subpart304-70.html). Publications derived from projects that were primary or ancillary to a cohort study funded by the NHLBI were also recognizable when they cited the parent study by name (eg, Multi-Ethnic Study of Atherosclerosis [MESA], Coronary Artery Risk Development in Young Adults Study [CARDIA]) but some did not. Those publications using data from an NHLBI study obtained through the NIH data repository BioLINCC (https://biolincc.nhlbi.nih.gov/home/) or dbGaP (http://www.ncbi.nlm.nih.gov/gap) were more challenging to identify if they did not acknowledge the data repository or, in the case of dbGaP, the specific study. Articles using National Health and Nutrition Examination Survey data not otherwise coded as NHLBI supported (see Table 2) were included if they used blood pressure or lipid data, because those components were specifically supported by NHLBI during this time interval. Publications were also considered NHLBI supported if they acknowledged relying on tools developed by NHLBI-supported studies or assistance by NHLBI staff, but these were few. Finally, publications that were silent on funding source were coded as unknown. After coding, the percentage of publications supported by NHLBI was calculated for each research domain. Results were summarized graphically to address question 1.

In addition, because the research domains are not mutually exclusive, the top 100 publications across all domains were pooled and duplicate publications were eliminated: the percentage with NHLBI support was calculated for the pooled but not duplicated publications.

To address question 2, the top 100 publications for each domain were broken into 3 successive 5-year intervals: 1998 to 2002, 2003 to 2007, and 2008 to 2012. The percentage of NHLBI-supported publications within each time interval was calculated based on the above methods, and the trends were graphed for each domain. The proportions of articles supported by NHLBI across the 3 time intervals for each domain were tested to determine whether they differed significantly by using a χ² test with 2 degrees of freedom. For those that were significant, a second test for linear trend was conducted with a χ² test with 1 degree of freedom. An additional analysis following a comparable process was conducted for the pooled unique publications across all domains.

For question 3, the top 100 publications for each domain were sorted by journal. The top 5 journals by frequency of top 100 publications for each domain were identified and results summarized.

Results

The total numbers of publications identified in the 10 domains over the 15-year period are summarized in Figure 1, using a logarithmic scale because the total number of publications across the selected domains varied by as much as 2 orders of magnitude. Fewest articles were published for the domain methodology with only 643 articles published over the 15-year period. The most articles were published for the domain risk factors with 86 017 articles published over the same time period. In total, >225 000 epidemiology articles were published in these 10 domains during the period 1998 to 2012.

Over the same period, 82.5% of the top 100 publications over the 10 domains provided sufficient information to code NHLBI support. Figure 2 presents the percentage of the top 100 publications within each research domain with NHLBI support based on average citations per year. Percentage of

---

Table 1. Domain-Specific Search Terms Used for SCOPUS Publications Search

<table>
<thead>
<tr>
<th>Behavioral Research</th>
<th>“adolescent behavior” OR “child behavior” OR “health behavior” OR “risk reduction behavior” OR “social behavior” OR “social psychology” OR “social adjustment”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comorbidity/syndromes</td>
<td>comorbid OR syndrome OR comorbid</td>
</tr>
<tr>
<td>Genetics</td>
<td>genetic* OR genomic* OR gene</td>
</tr>
<tr>
<td>Health disparities</td>
<td>“health status disparities” OR “healthcare disparities” OR “social class” OR “socioeconomic status” OR “socioeconomic factors” OR “factors,socioeconomic” OR “minority health”</td>
</tr>
<tr>
<td>Methodology</td>
<td>epidemiology AND methodology</td>
</tr>
<tr>
<td>New biology</td>
<td>proteomic* OR metabolomic* OR microbiome* OR epigenetic* OR RNA OR imaging OR exosome* OR “systems biology”</td>
</tr>
<tr>
<td>Preclinical disease</td>
<td>“preclinical OR preclinical OR subclinical OR subclinical”</td>
</tr>
<tr>
<td>Quality of care/clinical epidemiology</td>
<td>“quality of care” OR “clinical epidemiology”</td>
</tr>
<tr>
<td>Risk factors</td>
<td>“risk factor”</td>
</tr>
<tr>
<td>Surveillance</td>
<td>surveillance OR prevalence OR incidence</td>
</tr>
<tr>
<td>Additional search terms</td>
<td>AND (cardiovascular OR coronary) AND (human* OR “public health” OR clinical OR population* OR cohort OR epidemiol*) AND NOT randomiz*</td>
</tr>
</tbody>
</table>

---
publications with NHLBI support varied from a low of 12% for the domain quality of care/clinical epidemiology to a high of 46% for risk factors. The median proportion of publications with NHLBI support over the 10 research domains was 25%. The mean percentage of NHLBI-supported publications was 26.8% across the 10 domains.

When all the publications were pooled and duplicates were eliminated, the 1000 publications for 10 domains were reduced to 845. The percentage of NHLBI-supported publications among nonduplicative publications was 25.6%.

Figure 3 presents trends in the percentage of publications with NHLBI support over the 3 successive 5-year intervals for each of the 10 research domains. Trends suggested overall increases in 5 domains and overall decreases in 5, but percentages were statistically significantly different among time periods for only 2 domains: genetics (P=0.039) and quality of care/clinical epidemiology (P=0.024). When tested for linear trend, only quality of care/clinical epidemiology showed a statistically significant increasing trend (P=0.028). Again, pooling the data and eliminating duplicates provided an overall measure of the trend for NHLBI support of high-impact cardiovascular epidemiology research articles and indicated no significant change over the 3 time periods (P=0.056).

Finally, Table 3 presents the 5 most-represented journals within the top 100 publications for each domain. Additionally, Table 3 shows how many of the top 100 articles are in those 5 journals, and how many of those are supported by NHLBI. Among all domains, the most frequently appearing journals (in order) are Circulation, JAMA, The New England Journal of Medicine, JACC, and JAMA Internal Medicine. Within each domain, the 5 most-represented journals published one-third or more of the top 100 publications in most domains and ≥50% of the top 100 publications in 4 domains: risk factors, genetics, comorbidity/syndromes, and new biology. High-impact articles were most concentrated in fewer journals for genetics and risk factors where only 28 and 23 journals, respectively, accounted for the top 100 published articles; methodology was the most diffuse where 71 journals accounted for the top 100 published articles. The proportion of NHLBI-supported articles in the 5 most-represented journals of each domain in comparison with the proportion in the top 100 articles of that domain suggests the NHLBI-supported articles are equally or more represented among the articles in the most-represented journals.

For this analysis, 133 articles were reviewed by ≥2 coders to determine NHLBI support because they appeared in multiple domains with different coders. This masked comparison enabled an estimate of coding errors. Analyses of these duplicates, counting all possible pairs when ≥2 coders provided ratings, indicated 95.6% agreement among coders on NHLBI support.

**Discussion**

This analysis indicates that, based on citation analysis, at least one-quarter of the high-impact cardiovascular epidemiology research articles of the past 15 years were supported, at least in part, by NHLBI. Although an expectation for that percentage is difficult to determine, the estimate of ≈25% overall is higher than might be expected given that the proportion of biomedical research and development expenditures by the public sector in the United States in 2012 was $48.9 billion or 18.2% of the $268.4 billion (in US dollars) spent worldwide, of which NIH represents ≈$30 billion or 11.2% and NHLBI represents a little >$3 billion or 1.1%.

The 15-year time period was chosen to ensure sufficient time to obtain stable estimates for statistics and to evaluate recent trends within domains. Trend analyses did not indicate substantial change in the impact of NHLBI support for most research domains over the 15-year period. Only genetics and quality of care/clinical epidemiology showed significant variation over the 3 intervals with only the latter showing a significant upward linear trend in the impact of NHLBI support.
There are several implications of these analyses. First, NHLBI-supported cardiovascular epidemiology research accounts for about one-quarter of the top 100 high-impact articles published over the past 15 years. Considering that cardiovascular epidemiology is supported by many other government agencies and foundations, both in the United States and abroad, the NHLBI support of about one-quarter of the most highly cited epidemiology research is significant and suggests it is using its resources wisely. Second, the impact of NHLBI-supported research in quality of care/clinical epidemiology started from a low base but has significantly increased over the 15-year period and may reflect growing emphasis within the NHLBI on...
This analysis has a number of strengths. It covers a 15-year period that is of sufficient length to obtain stable estimates of the contribution of NHLBI support to high-impact cardiovascular epidemiology research. It includes 10 domains that offer multiple perspectives on the contribution of NHLBI-supported cardiovascular epidemiology research. It was conducted with few restrictions on the type of research support provided.

The analysis also has a number of limitations. An assessment of the impact of NHLBI-supported epidemiology research is not straightforward and, by necessity, must be captured indirectly. Attempts were made to capture most, if not all, the highest-impact publications in cardiovascular epidemiology over the period of study and to identify those with NHLBI support. Impact was measured by average annual citations that favor more recent publications over earlier publications, but the currency of the data was considered a higher priority than consistency of the follow-up period. The list of journals where high-impact articles were published favored those that have existed over the entire time interval examined and this may change with time because many new journals were introduced during the 1998 to 2012 time period. Unknown support information for 16.9% of publications may have altered the estimates of NHLBI support, although the estimates were conservatively calculated as if all support information were captured indirectly. Attempts were made to capture most, if not all, the highest-impact publications in cardiovascular epidemiology over the period of study and to identify those with NHLBI support. Impact was measured by average annual citations that favor more recent publications over earlier publications, but the currency of the data was considered a higher priority than consistency of the follow-up period. The list of journals where high-impact articles were published favored those that have existed over the entire time interval examined and this may change with time because many new journals were introduced during the 1998 to 2012 time period. Unknown support information for 16.9% of publications may have altered the estimates of NHLBI support, although the estimates were conservatively calculated as if all support information were known; therefore, the true impact of NHLBI-supported cardiovascular epidemiology research may be even greater. Finally, self-citations were not excluded from this examination. To investigate this effect, 2 domains were evaluated for the impact of self-citation, again using SCOPUS, and, at most, 8% of the articles changed, all within the bottom quintile of the top 100 articles. NHLBI support changed, at most, by 1%. Authors analyzed multiple perspectives on the contribution of NHLBI-supported cardiovascular epidemiology research. It was conducted with few restrictions on the type of research support provided.

The analysis also has a number of limitations. An assessment of the impact of NHLBI-supported epidemiology research is not straightforward and, by necessity, must be captured indirectly. Attempts were made to capture most, if not all, the highest-impact publications in cardiovascular epidemiology over the period of study and to identify those with NHLBI support. Impact was measured by average annual citations that favor more recent publications over earlier publications, but the currency of the data was considered a higher priority than consistency of the follow-up period. The list of journals where high-impact articles were published favored those that have existed over the entire time interval examined and this may change with time because many new journals were introduced during the 1998 to 2012 time period. Unknown support information for 16.9% of publications may have altered the estimates of NHLBI support, although the estimates were conservatively calculated as if all support information were known; therefore, the true impact of NHLBI-supported cardiovascular epidemiology research may be even greater. Finally, self-citations were not excluded from this examination. To investigate this effect, 2 domains were evaluated for the impact of self-citation, again using SCOPUS, and, at most, 8% of the articles changed, all within the bottom quintile of the top 100 articles. NHLBI support changed, at most, by 1%. Authors analyzed multiple perspectives on the contribution of NHLBI-supported cardiovascular epidemiology research. It was conducted with few restrictions on the type of research support provided.

The analysis also has a number of limitations. An assessment of the impact of NHLBI-supported epidemiology research is not straightforward and, by necessity, must be captured indirectly. Attempts were made to capture most, if not all, the highest-impact publications in cardiovascular epidemiology over the period of study and to identify those with NHLBI support. Impact was measured by average annual citations that favor more recent publications over earlier publications, but the currency of the data was considered a higher priority than consistency of the follow-up period. The list of journals where high-impact articles were published favored those that have existed over the entire time interval examined and this may change with time because many new journals were introduced during the 1998 to 2012 time period. Unknown support information for 16.9% of publications may have altered the estimates of NHLBI support, although the estimates were conservatively calculated as if all support information were known; therefore, the true impact of NHLBI-supported cardiovascular epidemiology research may be even greater. Finally, self-citations were not excluded from this examination. To investigate this effect, 2 domains were evaluated for the impact of self-citation, again using SCOPUS, and, at most, 8% of the articles changed, all within the bottom quintile of the top 100 articles. NHLBI support changed, at most, by 1%.
highest-impact cardiovascular epidemiology research publications, as measured by citation analysis, involved support, at least in part, by the NHLBI. Although expectations are difficult to calculate, the proportion of high-impact epidemiology research supported by the NHLBI is greater than budgetary data would suggest. These findings will inform future directions for NHLBI-supported epidemiology research and provide a baseline for future analyses to judge the comparative effects of any change in funding priorities at the NHLBI.

Acknowledgments
The authors thank David Gordon for insightful comments in planning the design of this article. Disclaimer: The conclusions of this article represent those of the authors and do not represent any official position of the National Heart, Lung, and Blood Institute; the National Institutes of Health; or the US Department of Health and Human Services.

Sources of Funding
This study was conducted without any grant support. The authors conducted this research as part of their salaried duties as full-time National Heart, Lung, and Blood Institute employees.

Disclosures
None.

References
Impact of National Heart, Lung, and Blood Institute–Supported Cardiovascular Epidemiology Research, 1998 to 2012

*Circulation.* 2015;132:2028-2033; originally published online October 27, 2015; doi: 10.1161/CIRCULATIONAHA.114.014147
*Circulation* is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2015 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circ.ahajournals.org/content/132/21/2028

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in *Circulation* can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to *Circulation* is online at:
http://circ.ahajournals.org//subscriptions/