Recently, popular magazines and newspapers broadcast skeptical headlines, such as, “Desperately Seeking Cures”; “A Decade Later, Genetic Map Yields Few Cures”; “Faltering Cancer Trials”; and “Grant System Leads Cancer Researchers to Play It Safe.” Some patients, physicians, advocacy groups, journalists, scholars, and policymakers openly wonder about the value of billions of dollars in government-supported research. During the past 10 years, the NIH budget has doubled, yet cardiovascular disease, while decreasing in incidence and severity, is still the leading cause of death and cancer incidence and death rates have declined little. Along with “Where Are the Cures?” critics ask “Who Is Accountable?”

Interest in Measuring the Performance of Science

There is a rapidly growing body of scholarship on the metrics of science, methods by which scientists, scientific organizations, policymakers, and funding agencies can gauge the value and impact of their work and investments. To assess the impact of the 2009 American Recovery and Reinvestment Act (ARRA), the National Science Foundation, the National Institutes of Health, and the White House Executive Office are engaged in an ambitious “STAR METRICS” project. Metrics often focus on publications and citations (“bibliometrics”) but also consider commercial products and quantitative impacts on practice or intellectual thought. Although scientists are typically eager to publish their work and see it cited, many worry about seeing their value to society, employers, and funders graded by black-box numbers. Some professional groups have harshly criticized bibliometrics, noting their inherent flaws and the risk that the act of measuring science will damage innovation, which is at the heart of the scientific enterprise.

At NHLBI, we increasingly recognize that we need to be held accountable for our performance, which means that it is incumbent on us to assess the performance of the scientists and projects we choose to support. It is a genuine challenge, however, for any nonprofit organization to assess its performance, because we cannot look to readily available metrics, such as revenues, profits, and stock prices. As in any business, it is important to align investments with expected returns. U.S. taxpayers are investing more than $30 billion a year in the NIH and expect more concrete measures than advancement in biomedical knowledge and a fiscal accounting of funds allocated to research. Many scientists whose career trajectories are largely dependent on securing NIH funding may be frustrated by critics who hold them responsible for high population burdens of disease, along with the inability of a fragmented healthcare system to exploit prior research discoveries. How can we manage expectations and encourage investments appropriate to desired outcomes?

A Thought Exercise About Unemployment and Jobs Training

Imagine that instead of funding research in cardiovascular biology and disease, we at NHLBI were asked to establish and run a jobs-training program in an impoverished city, a city that has seen its manufacturing base disappear and has not yet realized the benefits of the new service, technology-oriented economy. Our ultimate “end-product” would be a city bustling with full-employment, with high levels of personal income and socioeconomic well-being. As a community, we’ll need to bring together numerous partners and stakeholders, including legislators, government executives and agencies, employers, labor unions, and schools and universities. We’ll need to develop a multiple-pronged strategy to deal with short- and long-term factors that contribute to unemployment; one part of that strategy may be to establish a jobs-training program, the program for which we now find ourselves responsible.

Should our jobs-training program be held solely responsible for the city’s unemployment rate? Somehow, that does not seem right, as there are many factors contributing to unemployment that go way beyond what we can possibly address with the resources available to us. However, there are performance measures we could “accept” as reasonable reflections of the quality of our work. These might include the number of unemployed workers we train and the number of courses we offer. But these measures are descriptive only of the resources we have put in place and do not tell us whether we provided the right training to the right people. We could look for other measures of impact, like how many (and what proportion) of our trainees secured employment, how many (and what proportion) secured high-paying jobs, and how many employers choose to come to us to find future

The opinions expressed in this NHLBI Page are not necessarily those of the editors or of the American Heart Association.

From the Office of the Director, Division of Cardiovascular Sciences, NHLBI, Bethesda, MD.

Correspondence to Michael S. Lauer, MD, Office of the Director, Division of Cardiovascular Sciences, NHLBI, 6701 Rockledge Drive, Room 8128, Bethesda, MD 20892. E-mail lauerm@nhlbi.nih.gov

(Circ Res. 2011;108:405-409.)

© 2011 American Heart Association, Inc.

Circulation Research is available at http://circres.ahajournals.org

DOI: 10.1161/RES.0b013e3182125662
<table>
<thead>
<tr>
<th>Case Unit of Accountability/Strategy</th>
<th>City Jobs Program</th>
<th>Research Funding Agency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population (for communities, states, nation)</td>
<td>City residents and external businesses that may be attracted to relocate</td>
<td>U.S. and global population</td>
</tr>
<tr>
<td>Results</td>
<td>Full employment, economic well-being</td>
<td>Nation free of cardiovascular disease</td>
</tr>
<tr>
<td>Indicators</td>
<td>Unemployment rate, annual personal income, job satisfaction ratings</td>
<td>Cardiovascular mortality rate</td>
</tr>
<tr>
<td></td>
<td>Identify causes of unemployment (manufacturing shifts, tax policies, inadequate education)</td>
<td>Myocardial infarction incidence and case fatality</td>
</tr>
<tr>
<td></td>
<td>Identify new opportunities for employment</td>
<td>Identify scientific opportunities for investigation</td>
</tr>
<tr>
<td></td>
<td>Identify partners (employers, state government, labor unions, schools and universities)</td>
<td>Identify partners (researchers, policymakers, physicians, schools and universities)</td>
</tr>
<tr>
<td></td>
<td>Establish programs (eg, job-training program)</td>
<td>Establish programs (eg, government funding of researchers)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Performance (for programs/agencies)</th>
<th>Identify customers</th>
<th>Identify services</th>
<th>Identify metrics (over time—Figure 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identify customers</td>
<td>Unemployed workers</td>
<td>Provide training of new skills</td>
<td>How much? Courses offered (eg, computer programming, basics of laboratory technique)</td>
</tr>
<tr>
<td></td>
<td>Potential employers</td>
<td></td>
<td>Number of students enrolled</td>
</tr>
<tr>
<td>Identify services</td>
<td></td>
<td></td>
<td>Grants funded</td>
</tr>
<tr>
<td>Identify metrics (over time—Figure 1)</td>
<td></td>
<td></td>
<td>Research funds disbursed</td>
</tr>
<tr>
<td>How much?</td>
<td></td>
<td></td>
<td>Applications received and reviewed</td>
</tr>
<tr>
<td>How well?</td>
<td>Faculty hired and retained</td>
<td>Students who completed courses and passed exams</td>
<td>Grant reviews processed and completed according to schedule</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Funds disbursed according to schedule</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Progress reports received, reviewed, and acted upon according to schedule</td>
</tr>
<tr>
<td>What impact?</td>
<td>Number or proportion of students who secured jobs</td>
<td>Average income of students who secured jobs</td>
<td>Clinical trials that recruit subjects and complete processes on time and within planned budgets</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Publications (including number, cost per publication, appearance in high-impact journals)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Citations (including number, cost per citation, appearance in major systematic reviews or guidelines, appearance in high-impact journals, h-index)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Commercial (including patents, products readied for clinical testing, cost per patent)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Economic (jobs created, additional business activity stimulated)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Practice or policy (including changes to guidelines or law, changes in evidence-base prescription practices)</td>
</tr>
<tr>
<td>Tell the stories behind the metrics</td>
<td>Are there measurement artifacts that convey misleading impressions?</td>
<td>Are there measurement artifacts that convey misleading impressions?</td>
<td>Are there measurement artifacts that convey misleading impressions?</td>
</tr>
<tr>
<td></td>
<td>Were the appropriate skills taught?</td>
<td>Is peer review identifying the best scientists and best projects?</td>
<td>Is competition from industry, foundations, and agencies from other countries crowding us out?</td>
</tr>
<tr>
<td></td>
<td>Were students correctly matched to courses?</td>
<td>Were budgets appropriate and are they adequately flexible? Do they account for dynamic changes in technologies and prices?</td>
<td>Are powerful economic forces keeping researchers and/or patients out of government-sponsored projects? Are economic forces preventing adoption and dissemination of evidence-based technologies?</td>
</tr>
<tr>
<td></td>
<td>Did we adequately research employer needs?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
employees. If we focus on these measures, we can demonstrate our value to the community, and equally important, we can make informed decisions about changes we can make that will improve our efficiency and effectiveness.

**Results-Based Accountability**

We have just applied an approach called “Results-Based Accountability (RBA),” which is “a disciplined way of thinking and taking action that can be used to improve the quality of life in communities, cities, counties, states, and nations. Results-based accountability can also be used to improve the performance of programs, agencies, and service systems.” During the past few years, a number of NIH leaders have learned about RBA as part of a formal NIH senior leadership program run in conjunction with the University of Maryland School of Public Policy. I summarize how we applied RBA to unemployment and job training in Table 1.

In RBA, we start by thinking about population ends, in this case an economically successful city operating at full employment. We then move to the role of our jobs-training program and come to realize that we should focus on means, in this case training unemployed workers and enabling them to secure new jobs. We identify measurable and meaningful performance metrics, such as the proportion of our trainees who secure jobs. RBA is a helpful construct because it forces us to separate population ends from program performance.

In Table 1 and Figure 1, I summarize a similar thought exercise about how we, a Federal funding agency that supports biomedical researchers, can use the RBA construct to evaluate our performance. We proceed through a series of steps as follows.

1) Identify our customers and articulate the services we provide (eg, provide funds to applicants who submit highly meritorious proposals).

2) Identify a relatively small number of “headline” performance metrics. Some metrics focus on process (eg, accruing patients to trials, completing application reviews on time, answering queries from applicants and grantees), whereas others focus on impact (eg, publications, citations, ancillary studies, new commercial products, changes in clinical practice).

3) For each metric, plot changes over time and forecast what is likely to happen if we make no changes in our processes.

4) “Tell the story” behind the metric. We ask a series of questions to identify factors that have improved or worsened performance.

5) Based on our story, identify partners and possible strategies to “turn the curve,” that is, attempt to change adverse trends.

**Performance Metrics for Biomedical Research**

To identify what performance metrics might be useful, we can do another thought experiment. Suppose we were an outright failure, that is, we funded universities and researchers who did absolutely nothing scientific with our money. How would we know? We would find no publications that could be linked to our grants. With no publications, there would be no citations. We might also find no evidence of new products being made ready for clinical testing and eventual commercialization (ie, no patents). We would see no trials going anywhere near completion, resulting in no publications, citations, or guidelines citing our research for recommendations on changes in practice.

This worst-case scenario is easy for all of us to assess. In the real world, it is more difficult to assess whether a given level of productivity is appropriate in quantity (numbers of
predictors of future scientific success and recognition and scholars have suggested that bibliometrics are powerful tools for assessing research output. They have led to at least 1267 R01 grants that were awarded between 1990 and 2009 through the NHLBI cardiovascular research division and that were classified as “nonhuman.” To identify these grants, I used the keyword “myocard*” to search for relevant publications in the PubMed database.

Using an internal NIH-based search tool, I identified at least 2.2 billion dollars in funding for NIH grants and their successful competitive renewals from 1990 to 2009. These grants have been widely cited, with at least 15,656 publications that have garnered 448,830 citations (approximately 29 citations per publication). The average cost per publication is $49,977, and the average cost per grant number is $143,276. These figures compare favorably to national estimates of cost per publication for all academic-based research, which has increased from $186,567 in 1998 to $308,641 in 2008.

Figure 2 shows the behavior of selected process and impact metrics over time. There has been a marked increase in the number of projects funded and total funding (even after accounting for inflation), coincident in part with the NIH doubling in the late 1990s and early 2000s. To allow for fair comparisons, the output metrics related to publications and citations are shown for each project from its start until 5 years later. The number of publications increased commensurate with the increase in funding, whereas the cost per publication remained remarkably constant. However, the number of citations has not kept pace, with a decrease in the number of citations per publication and an increase in cost per citation.

Our next step, which is critical, is to ask a series of questions that will reveal the story behind these data. This is an exercise that is beyond the scope of this article. We might start by asking why the citation rate is going down. Is it because the field is losing luster? Or, perhaps, is it because we account for the number of citations per publication and an increase in cost per citation.

An Incomplete Example

Using an internal NIH-based search tool, I used the keyword “myocard*” to identify 1267 R01 grants that were awarded between 1990 and 2009 through the NHLBI cardiovascular research division and that were classified as “nonhuman.” To date, these grants and their successful competitive renewals account for $2.2 billion in funding. They have led to at least 15,656 publications that have garnered 448,830 citations (approximately 29 citations per publication). The average cost per publication is $49,977, and the average cost per grant number is $143,276. These figures compare favorably to national estimates of cost per publication for all academic-based research, which has increased from $186,567 in 1998 to $308,641 in 2008.

Figure 2 shows the behavior of selected process and impact metrics over time. There has been a marked increase in the number of projects funded and total funding (even after accounting for inflation), coincident in part with the NIH doubling in the late 1990s and early 2000s. To allow for fair comparisons, the output metrics related to publications and citations are shown for each project from its start until 5 years later. The number of publications increased commensurate with the increase in funding, whereas the cost per publication remained remarkably constant. However, the number of citations has not kept pace, with a decrease in the number of citations per publication and an increase in cost per citation.

Our next step, which is critical, is to ask a series of questions that will reveal the story behind these data. This is an exercise that is beyond the scope of this article. We might start by asking why the citation rate is going down. Is it because the field is losing luster? Or, perhaps, is it because we funded a number of relatively new investigators during the doubling, whose projects have not yet matured to the point at which they can garner many citations?
Closing Thoughts

As stewards of taxpayer monies, we at NHLBI expect to be held accountable for our decisions and policies. No differently than scientists, hospital administrators, and for-profit company executives, we look to performance metrics as tools—nothing more or nothing less—to help understand the stories behind the effectiveness of our work and to help us make the best informed decisions. We look to colleagues in NIH, in the extramural scientific community, in medical editing, and in professional societies to work with us as we develop greater degrees of sophistication in program evaluation, ultimately to reach the goal of a society free of the ravages of cardiovascular disease.

References


Thought Exercises on Accountability and Performance Measures at the National Heart, Lung, and Blood Institute (NHLBI): An Invited Commentary for Circulation Research
Michael S. Lauer

Circ Res. 2011;108:405-409
doi: 10.1161/RES.0b013e3182125662

Circulation Research is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2011 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7330. Online ISSN: 1524-4571

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circres.ahajournals.org/content/108/4/405

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation Research 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 Research is online at:
http://circres.ahajournals.org/subscriptions/