# Chapter 3
# Training for Informatics Research Careers: History of Extramural Informatics Training at the National Library of Medicine

**Valerie Florance**

**Abstract** The National Library of Medicine (NLM) has been the primary funder for university-based informatics training programs in the U.S. since the early 1970s. NLM has provided institutional training grants as well as informatics research opportunities for individual fellows. The programs supported by NLM have changed over time as the competencies needed for informatics research training have evolved. Over the years the focus of the program has broadened to address a wide range of informatics needs, including the incorporation of bioinformatics and public health informatics training into programs that had earlier been focused almost exclusively on medical informatics. This chapter describes the evolution of grant-supported informatics training, identifies basic elements of informatics curricula designed to produce informatics researchers, highlights best practices in program administration, and discusses models for program evaluation that can be applied to the informatics training programs.

For 40 years, the National Library of Medicine (NLM) has been a major source of federal support for university-based training in biomedical informatics. NLM received its authority for providing training through the Medical Library Assistance Act (MLAA). Signed into law in October 1965, MLAA authorized NLM to train librarians and other information specialists. Between 1965 and 1970, about 11 % of NLM's grant budget was spent to support training [1]. By comparison in 2012, 23 % of NLM's grant budget was spent to support informatics research training. In 1971, "training for biomedical communications careers… included Master's degree programs in library science, and doctorate programs in health information research and the history of medicine. In addition, there are post-doctoral research fellowships and library internships for advanced training in information processing and medical

V. Florance, PhD
National Library of Medicine, NIH, DHHS, 6705 Rockledge Drive,
Rockledge 1, Suite 301, Bethesda 20892, MD, USA
e-mail: florancv@mail.nih.gov

E.S. Berner (ed.), *Informatics Education in Healthcare*, Health Informatics,
DOI 10.1007/978-1-4471-4078-8_3, © Springer-Verlag London 2014

librarianship" [2]. In 2012, NLM support for informatics research training was confined to predoctoral and postdoctoral work at its university-based programs. NLM's predoctoral trainees are expected to obtain PhD degrees, and most NLM postdoctoral fellows receive an MS or PhD degree.

In the early 1970s, believing that the shortage of health sciences librarians that had led to its original training authority was no longer as severe, NLM redirected the focus of its training programs from librarian training to health scientist training in the use of computers in medical research, education and healthcare (p. 407) [3]. According to NIH grant records, in 1972 there were NLM-supported informatics training programs at Duke, Stanford, UCSF, and the University of Alabama-Birmingham, plus librarian or biomedical communication training programs at Case Western Reserve and Georgia Institute of Technology. In 1973, the Department of Health, Education and Welfare (now known as the Department of Health and Human Services) directed its agencies to end federal support for biomedical training, which reduced the flow of support to informatics training for several years. The 1978 NLM Annual Report indicates that NLM-supported training programs were now located at ten institutions:: University of California, San Francisco (UCSF), University of Alabama-Birmingham, Duke University, Ohio State University, University of Minnesota-Minneapolis, Case Western Reserve University, University of Missouri-Columbia, Mt Sinai School of Medicine, Georgia Institute of Technology and the University of Illinois Urbana-Champaign. These programs supported 71 trainees. Of the 41 predoctoral trainees, 16 sought Masters degrees and 23 sought PhDs. Two thirds of the 30 postdoctoral trainees were physicians (MDs) (pp. 50–51) [4]. The number and location of the universities providing informatics training have changed across the years, though a core group of universities has provided NLM-funded informatics research training for more than 20 years. Changes were caused by budget fluctuations and by a changing view of the programs' purposes. In the 1970s, NLM trainees were mostly physicians learning to use computers to manage health information. But in 1983, NLM saw a need to focus on research career training and a new funding announcement was issued (p. 34) [5]. As noted in the 1984 report "Research issues in the health information and health computer sciences call for highly trained, creative talent, able to articulate medicine with computers and healthcare with information science. There is a particular need in academic medicine for a new discipline – health information or health computer science. Through its training program, NLM provides grants for research career training in this field of medical informatics (p. 38)" [6]. At this time, NLM articulated the basic components of career training in informatics: … "didactic instruction, involvement in major, ongoing health computer science studies; and opportunities for work in advanced information science research" (p. 38) [6]. Five programs received funding as a result of this new offering: UCSF, Minnesota, Harvard, Tufts-New England Medical Center and Stanford. Over the ensuing years, NLM supported as many as 18 separate programs. Most recently, as of July 1, 2012, 14 programs received new 5-year awards for training of 108 predoctoral, 79 postdoctoral and 31 short-term trainees, plus 9 pre- and postdoctoral dental informatics trainees funded by the National Institute of Dental and Craniofacial Research [7].

Two of these programs were from the group of five funded in 1984 and seven of the current NLM programs have been providing NLM-supported informatics training for more than 20 years.

## Evolving Scope of NLM's Informatics Training Programs

Beginning in 1996, NLM's funding announcements provided flexibility for applicants to suggest specialized training possibilities, including special support for librarians, cancer informatics and dental informatics (in the latter cases, other NIH Institutes provided funds for specialist training at NLM-funded programs). In 2001, in the midst of the doubling of the NIH budget, NLM expanded the scope of informatics training, noting that "NLM is aware that informatics has historically had a heavy focus on clinically relevant topics, and that healthcare delivery continues to offer a rich variety of important research questions for informaticians. However, the remarkable emergence of very large datasets in genomics, neuroscience, clinical research, health services research and other domains has resulted in a rapidly expanding interest among basic and clinical scientists in the potential of informatics for facilitating research and for creating knowledge. NLM believes there will be high demand for specialists capable of applying informatics to biomedical research. Core training for informaticians should include exposure to the informatics of biomedical research" [8]. This emphasis of NLM led to an infusion of bioinformatics into what had been solely clinical informatics programs (see also Chap. 11). As a result, many informatics programs added new faculty who could teach this material.

In the 2006 funding announcement, applicants were strongly encouraged to require a degree from most trainees, including postdoctoral fellows. Four thematic training domains were proposed, healthcare/clinical informatics; bioinformatics and/or computational biology, clinical research and translational informatics, and public health informatics. In addition to addressing at least one of these areas, applicants could propose specialized tracks in education of health professionals, imaging and signal processing, health services research, or another area if pre-approved by NLM. Most of the successful programs offered at least two of the main domains, such as healthcare/clinical and public health, while some offered all four. Few proposed specialized tracks [9].

In the most recent solicitation, issued in 2011, NLM used the four informatics tracks used by the American Medical Informatics Association (AMIA) as the core areas for informatics research training, developing a brief definition for each:

- Healthcare/clinical informatics (HCI): Applications of informatics principles and methods to direct patient care, such as advanced clinical decision support systems and multimedia electronic health records, to the provision of informational support to healthcare consumers. Special tracks might be offered for nursing informatics, dental informatics, imaging informatics, or other appropriate clinical themes.

- Translational bioinformatics (TBI): Applications of informatics principles and methods to support 'bench to bedside to practice' translational research, such as genome-phenome relationships, pharmacogenomics, or personalized medicine. Special tracks might be offered in health effects of environmental factors, genome-wide association studies (GWAS), or other similar areas.
- Clinical research informatics (CRI): Applications of informatics principles and methods to support basic clinical trials and comparative effectiveness research. Special tracks might be offered in areas such as biostatistics, in-silico trials, merging and mining large disparate data sets that mix images, text and data.
- Public health informatics (PHI): Applications of informatics principles and methods to build integrated resources for health services research, for decision support in public health agencies, to support regional or global health research, or syndromic surveillance. Special tracks might be offered in areas such as health literacy, information design for consumers, health effects of climate change [10].

## Using the NIH Guide to Track the Evolution of Informatics Research Training

The NIH Guide to Grants and Contracts is a regular issuance of all new funding announcements offered by the National Institutes of Health. Announcements of new competitions for NLM-funded university-based research training are published there, every five years. In a sense, the funding opportunity announcements published in the NIH Guide to Grants and Contracts provide snapshots of the evolution of informatics research training supported by the National Library of Medicine. For example, in 1996, the purpose was stated this way: "Such training will help meet a growing need for qualified, talented investigators, well prepared to address information problems in healthcare, health profession's education, biomedical research, health services research, and public health" [11]. The 2001 issuance states it this way: "Graduates of the NLM-supported programs should be able to conduct basic or applied research at the intersection of biology and medicine with computer and cognitive sciences, and are expected to be familiar with the use and potential of modern information technology" [8]. In 2006, applicants were told: "Such programs help meet a growing need for investigators trained in biomedical computing and relevant application domains including healthcare delivery, basic biomedical research, clinical and translational research, public health, health information sciences and other related areas. Informatics training is multi-disciplinary. This initiative is not intended to prepare trainees for careers emphasizing planning, deployment, maintenance, or administration of computer systems in healthcare, public health, medical education or research. The emphasis in this program is on the development of new knowledge that advances informatics as a scientific discipline" [9].

In the most recent announcement for NLM training programs, issued in 2011, the purpose statement noted that: "Graduates of the NLM-supported programs should

be able to conduct original basic or applied research at the intersection of computer and information sciences with one or more biomedical application domains. Successful graduates of these programs will be prepared for research-oriented roles in academic institutions, not-for-profit research institutes, governmental and public health agencies, pharmaceutical and software companies, and healthcare organizations. This initiative is not intended to prepare trainees for careers emphasizing planning, deployment, maintenance, or administration of computer systems in healthcare, public health, medical education or research. The emphasis in this program is on the development of new knowledge that advances informatics as a scientific discipline" [10].

NLM's funding announcements document the scope of informatics training at points in time. But in the years between grant competitions, NLM sometimes expanded the scope of training by awarding grant supplements to the existing programs. Following the issuance in 1999 of the Biomedical Information Science and Technology Initiative (BISTI) report [12], NLM awarded supplemental funds in 2000 and 2001, to its existing training programs, to support development of resources for training bioinformaticians. In these years, supplements were also awarded to NLM training programs to strengthen offerings relating to health services research (pp. 68–69) [13]. In 2005, the Robert Wood Johnson Foundation awarded a grant to NLM to support the development of public health informatics as a research career (p. 59) [14]. As a result, four NLM training programs received supplemental support to fulfill the objectives of this initiative, and several that did not receive funds implemented tracks for public health informatics.

In summary, the field of informatics has evolved gradually over decades, as observed through the prism of NLM's informatics training programs. From the idea that computers could help physicians with health information management, informatics as a scientific domain now has multiple subfields of interest ranging from clinician decision support to computational modeling of disease processes to global monitoring of disease outbreaks to patient-controlled health records to information interfaces for low literacy populations.

Although university-based programs have been NLM's core strategy for training a cadre of informatics researchers, at times since 1972, NLM has also supported extramural informatics training for individuals, and a number of other NIH institutes employ individual fellowships as a mechanism for research training. Between 1989 and 2005, NLM awarded dozens of individual fellowships for research or applied informatics to individuals not enrolled at one of NLM's university-based programs. For example, in 1992 NLM announced an individual applied informatics fellowship program, noting in the announcement that "If informatics is to realize its full potential as an indispensable tool for researchers and health-care workers, there must be adequate number of health professionals able to apply the knowledge of informatics to develop modern information systems in traditional organizations, use the new information techniques in a specific field, and help disseminate promising programs and systems" [15]. Although NLM does not currently offer individual fellowships, the NIH Advisory Committee to the Director, Biomedical Workforce Task Force [16] recently recommended that all components of NIH offer individual predoctoral fellowships for research training. Implementation of this

recommendation, planned for 2014 and 2015, will provide expanded access to individual fellowships for informatics training.

While NLM has been the primary source at NIH for informatics funding, other Institutes have training grant programs or fellowships that encompass informatics elements focused in a particular domain. For example, the National Institute of General Medical Sciences offers predoctoral training in bioinformatics and computational biology and in biostatistics. The National Institute of Biomedical Imaging and Bioengineering offers training in Imaging and Information Sciences. The National Cancer Institute offers individual fellowships which support research in a clear cancer focus area. The NIH Guide [17] and the home web sites of the 24 NIH Institutes and Centers that make awards [18] provide greater detail on the scope and focus of such informatics-related training.

## Data Requirements for Training Grant Applications

Every application to NIH for training funds, whether from NLM or another Institute, requires a detailed data set reflecting institutional resources and readiness to train top quality biomedical researchers. Three elements of any Funding Opportunity Announcement (FOA) or Request for Applications (RFA) related to institutional training provide the interested applicant with insight into the critical content elements.

### *Description of the Funding Opportunity*

This section gives the basic outline of the type of training program, levels of trainee to be supported, the offeror's rationale for investing in training and, sometimes, what is not covered by the offering. In NLM's most recent offering (RFA-LM-11-001), the description section provides guidance on the different areas of informatics training that could be proposed, the fundamental elements of the curriculum, (e.g., core curriculum plus a range of electives); the support, both technical and human a trainee should receive (e.g., meaningful, supervised research experience); and the intended product of training (e.g., independent research compatible with publication of results and competition for grants). Expected endpoints of training are also listed, as are options for specialized training themes or tracks [10].

### *Research Program Plan*

In RFA-LM-11-001, this section provides a more detailed picture of expectations regarding program administration (e.g., administrative home of the program); faculty (both core and collaborating faculty); and proposed training (e.g., long term

objectives of the program and strategies for carrying them out. Specifications include details about core curriculum, practicum experience, elective options and trainee research experience); evaluation plan for the program as a whole and for individual trainees; pool of candidates; institutional environment, including an estimate of other similar trainees at the institution. Reviewers scrutinize this section for evidence that applicants have thought out the details, have a strong curriculum and rich environment with collaborating faculty from other departments, and have past success at training. Reviewer analysis of the research program plan is done in conjunction with the data tables required in all training grant applications (Tables 1–10 for new applications, Tables 1–12 for renewals) [10].

## *Data Tables for Training Grant Applications*

All NIH training grant applications use the same data tables to provide reviewers and grant program staff with evidence about the program's past success and/or likelihood of future success. The tables collect detailed evidence in the following areas: participating departments/programs and faculty; other institutional training grants in the participating units; grant support of participating faculty; pre- and post-doctoral trainees of participating faculty; publications by pre- and post-doctoral trainees; admissions and completion records for participating organizations; qualifications of recent and current applicants; admissions and retention of underrepresented populations. Previously funded programs must also submit tables covering pre- and post-doctoral trainees supported and their current status. Careful thought should be given to these tables in light of the proposed plan. For example, thoughtful selection of collaborating departments and faculty could strengthen an application. Reviewers have a keen eye for data that do not resonate with the textual content. NIH provides an extensive set of templates and instructions for these tables. An application missing these data will not review well, so devoting time to gathering and reviewing tabular information in advance is a wise investment.

## *Scored Review Criteria*

All grant solicitations, training or otherwise, include lists of scored criteria and additional review criteria. The former affect the overall impact score, the latter do not with one exception. Although it seems obvious, the importance of addressing scored review criteria within the body of the application cannot be overstated. In RFA LM-11-001, review criteria are listed as questions. An applicant should know where the answers are to these questions in his/her application. Having an outside reader try to answer them might be a useful pre-submission exercise, to assure that all points are well-covered. Among the additional review criteria, most are not applicable to training grants. However, if there is a section called "Renewals",

applicants who are seeking a new round of funding for their existing training grant should be certain that these questions are answered in addition to those listed in the scored review criteria section. These added questions relate to how well the applicant performed in the past funding period. Reviewers look carefully at renewal applications and will assign poorer scores to those that do not show strong results.

## Model Training Program for Biomedical Informatics

While NLM has never dictated the specifics of curriculum content or program structure for the research training it supports, funding competition announcements have always enumerated the important factors to be incorporated into a research training plan. These always include:

1. Interdisciplinary content, with coverage of information science, cognitive science and knowledge of one or more domains of biomedicine
2. A core curriculum of required courses emphasizing informatics concepts and methods and state-of-the-art technology assessments
3. Electives providing opportunities for advanced training in informatics fields
4. Individual research experience for each trainee, including assistance for trainees in selecting appropriate research projects
5. Exposure to the informatics of basic biomedical research
6. Effective programs for recruiting and retaining a diverse pool of trainees
7. Approaches for evaluating program success

During a project to develop and revise NLM's overall training program evaluation framework, data were extracted from more than a dozen training programs over 15 years of training experience. Analysis yielded patterns of activity that characterize successful programs [19]. They are framed below as four program objectives for a model training program.

***1. Produce researchers prepared to conduct independent research in biomedical informatics by the time they complete their training.***

Key to a successful training program in informatics is attracting and retaining a diverse group of trainees. For some programs, trainees are selected from a pool of outside applicants. For others, they are selected from the University's matriculated graduate student population (the latter approach is most common in biological sciences areas such as bioinformatics or computational biology). Although some fields assume that new PhDs will obtain postdoctoral training, this is not always the case in biomedical informatics. For example, a study of NLM trainees who graduated between 1991 and 2005 showed that about 15 % continued their training through postdoctoral appointments, residencies or additional graduate degrees (Table 42, p. 65) [20]. An evaluation metric for this area might be that 95 % of graduates have obtained a suitable position, a career transition award or entered postdoctoral training within 1 year of completing planned training.

Outside funders usually have time limits for support of graduate training. For example, NLM provides up to five total years of support for predoctoral training, or three years of postdoctoral support. A recent report of an NIH Biomedical Workforce Task Force recommends no more than five years of total support for graduate training [16], a reduction of what has been allowed in past years. While all universities have multiple sources of support for their graduate students, good management practice suggests that programs which depend heavily on outside funding should be setting some numeric targets within the institution, such as 90 % of predoctoral trainees complete their planned training within five years.

***2. Provide state-of-the-art informatics curriculum content, successful research mentors, research practicum opportunities to a diverse group of trainees during the training period.***

This objective involves multiple tasks and targets. Establishing a continuous program of curriculum review and renewal means setting a threshold, such as 25 % of courses are refreshed each year, or one new course every two years. A plan for dropping or replacing courses would be part of this process.

In its funding solicitation of 2006, NLM characterized a core curriculum as curriculum "addressing informatics concepts and methods that support the entire program, spanning all application domains that are addressed….the preponderance of courses and other educational elements comprising the core must apply to all application domains." Applicants were required to provide details about component courses and educational experiences [9].

All of NLM's university-based training programs offer a core curriculum of required courses plus an extensive menu of elective courses, often housed in collaborating departments such as computer science or business or molecular biology. All offer at least one core course in the basic principles and concepts of informatics. For the university-based training programs funded by NLM in 2006, the most common required courses (in addition to the core course) include quantitative methods (69 %) and techniques of computer science, engineering or other information fields (75 %), biological sciences (50 %), ethics (50 %) and research methods (50 %) [20]. When prerequisites are required, as they are at several programs, they are typically courses in computing or quantitative methods. Many programs require more than one core course, so that the subfields of informatics can be covered in greater detail. The number and scope of required courses reflects the philosophy of the program director – some programs are tightly structured into tracks, while other allow a trainee to tailor the coursework plan for research area that interests her/him.

Engaging faculty in a way that advances their own work as well as those of the trainees requires action at several levels. New faculty should receive some kind of training in mentoring and/or have mentors of their own. If a target is set that 100 % of trainees in translational bioinformatics have access to dual mentors, then the dual mentors will need to learn this role. Targets should be established for faculty publishing and research activities such as 90 % of faculty have outside funding for their research and peer-reviewed publications in their research area. Programs should provide mentorship training to faculty who will be mentoring the trainees.

To establish standards for increasing diversity in the trainee pool, targets might be established for each type of intervention, such as to attend two minority-focused recruitment meetings per year or offer at least three short-term research experiences for underrepresented groups. NIH requires all training programs to offer in-person training for Responsible Conduct of Research (RCR) to 100 % of their trainees. Many universities developed online learning for this training, but NIH requires in-person RCR training as well, so targets must be set for in-person experiences too, such as introducing a case study that fulfills RCR requirement into all core courses.

A recent report of the Biomedical Task Force of the NIH Advisory Committee to the Director reported that NIH trainees were occupying an array of science-related positions, rather than solely occupying academic-style tenure track positions. They recommended that "NIH should create a program to supplement training grants … to allow institutions to provide additional training and career development experiences to equip students for various career options" (p. 8) [16].

In the past, the expectation for a biomedical research trainee supported by NIH was that she/he would graduate, obtain an academic position, and begin to apply for research project grants from NIH. The Task Force report acknowledges that only about 43 % of NIH research trainees follow that path, and asserts that science-related careers, in government, in industry and public policy, can be as important to the advancement of science as academic pursuits (see Figure 19, p. 32) [16].

Studying its own trainees in 2008, NLM found that about 40 % held faculty positions, 21 % were working in industry or self-employed in small businesses, 16 % worked in healthcare organizations, 15 % were still in training, and the rest worked in government agencies or other non-profit organizations (Table 42, p. 65) [20]. Additionally, a pattern emerged suggesting that across their careers, informaticians often move back and forth among these options. An important lesson to be drawn here is that curriculum planners should think about the underpinnings of a research career broadly, and prepare trainees for administrative and managerial roles as well as for research. For example, in 2010, when NLM provided curriculum development funds to its university-training programs, the program at Rice University created an online course covering topics such a lab management and grant writing [21], and several other programs offer electives in these areas.

***3. Advance knowledge in the field of biomedical informatics during and after the award period.***

Funding agencies are increasingly focused on measuring the outcomes and impact of the grants they award. The area of advancing knowledge is often measured bibliometrically, using publication and citation rates. Bibliometric methods are limited in their ability to capture the full range of informatics trainee contributions. Analysis of 200 peer-reviewed articles published in 2012 by NLM grantees shows that 34 % of the articles cite training grant numbers, but several programs not represented indicated that their trainees had, indeed, published articles. One reason for this is that peer-reviewed publications or other dissemination venues that don't include the grant number in an acknowledgement are difficult to identify. Another is that commercial resources available for bibliometric analysis, such as Scopus or Web of Science or Google Scholar don't cover the full range of journals in which informaticians publish. Nevertheless, training program directors can set targets for

this area such as 95 % of trainees author or co-author a peer reviewed article that is published or accepted for publication and 100 % of trainees make at least one presentation of their work at a national meeting.

Software, datasets or knowledge resources produced by informatics researchers are not easily identified by bibliometric methods, and there are no widely accepted metrics for these types of contributions. Implementation planning at NIH for recommendations from the Advisory Committee to the Director's Working Group on Data and Informatics includes developing approaches for identifying and citing datasets and knowledge resources that could be important to informaticians [22]. When such resources are catalogued, it will be easier for trainees to cite their contributions and for programs and funders to monitor performance.

***4. Demonstrate administrative competence through program management and evaluation.***

Every training program must undertake regular evaluation of the entire program as a whole. In NLM's current group of training programs, two types of internally sponsored program evaluation are most common. Several programs have an external advisory group that meets every year and provide advice to them. In some universities, the graduate school has a regular five year cycle in which the university brings together a committee to evaluate the program. Peer review of a training grant proposal can provide valuable extramural assessment of proposed curriculum and activities.

In addition to overall evaluation, training programs must establish evaluation metrics for each type of training activity. Approaches taken for establishing trainee evaluation metrics vary by university, but each NLM-funded program employs such metrics, which often involve course completion, academic achievement, publications, presentations, awards and evidence of leadership. Some programs employ explicit core competencies for each curriculum component. The recently-issued AMIA Academic Forum report provides an excellent starting place for a university considering a program in biomedical informatics training, providing both a definition of biomedical informatics and a set of competencies to drive core curriculum [23].

Establishing a personal training plan for each trainee and providing career counseling and other resources to assist them in the transition from training to career are fundamental activities of a model program. The Biomedical Workforce Task Force Report noted a lack of consistency in the mentoring provided in the training models supported by most individual NIH Institutes. Training supported by other NIH Institutes falls under the National Research Service Award (NRSA) rules. For a typical NRSA predoctoral trainee in molecular biology, two years of predoctoral support through a T32 training grant would be followed by several additional years supported as a graduate assistant paid by the research grant(s) of a mentor/investigator. For trainees in this model, the Task Force felt that individual development plans, career counseling and tracking of trainee accomplishments needed strengthening (pp. 8–9) [16]. NLM's training programs, which are not part of NRSA, employ a different strategy for trainee support, one that does not have these deficiencies. NLM provides funding for up to five years of predoctoral training through the training program itself, so that trainees work with their mentors over an extended period. Typically, trainee progress is evaluated twice each year, in writing, by the faculty

mentor, based on a training plan worked out at the beginning of a trainee's learning program.

Programs must also have methods to track trainee accomplishments and make that information available to prospective trainees. Another planned NIH initiative based on recommendations in Reports of the Biomedical Workforce Task Force and the Working Group on Diversity in the Biomedical Research Workforce involves developing a tracking system for trainees who have been supported by federal funds, so that long term career development and publication patterns can be analyzed, and future grant applications can be simplified by pre-filling certain fields [16, 24]. It is felt that this will deliver value to both the funder and the training organization, as universities can use this information to recruit future trainees. To date, there is no single system used in universities to track students in this way. CareerTrac, a tracking system developed by the National Institute of Environmental Health Sciences and the Fogarty International Center, is being implemented by NLM's informatics training programs in 2013. CareerTrac [25] allows easy linking of trainees with papers listed in Pubmed and training information in their appointment forms, and allows tracking of awards, presentations and career steps. CareerTrac, or a system based upon it, will likely be implemented for all NIH training programs in the next three years.

Recruitment and retention of a qualified, diverse group of trainees has always been a criterion for success listed in NLM's grant funding announcements for informatics training. The 2011 solicitation stated: "This FOA requires that all applicants submit a diversity recruitment and retention plan. While applicants may base their plans on multi-disciplinary programs in place at their institutions, they must also indicate how the informatics programs will participate in these recruitment activities and how these activities will meet the needs of potential applicants with interests in informatics" [10]. Reporting on past success at minority recruitment is a requirement of all NIH training grant applications (Table 10). Among NLM's programs, strategies include offering special summer programs; attendance at regional and national meetings such as the American Indian Science and Engineering Society or the Annual Biomedical Research Conference for Minority Students; partnership with one or more minority-serving universities, tribal colleges or historically black colleges and universities.

## Evaluation Framework for NLM's Extramural Training Program in Biomedical Informatics

In 2007, working with Humanitas, a management and technology consultant, NLM grant program staff developed a framework for overall evaluation of the informatics research training programs sponsored by NLM [20] (Appendix B). Beginning with the program goals stated in NLM's early funding announcements, three training program objectives were framed, standards and indexes were established for each

objective, and data points identified that would indicate the level of success. The three initial objectives were: Increase the number of researchers trained to conduct independent research in the field of biomedical informatics; Develop and increase institutional training capacity for the field of biomedical informatics; Advance knowledge in the field of biomedical informatics. The example below shows the development of this concept for one objective.

Objective 1: Increase the number of researchers trained to conduct independent research in the biomedical informatics field during the award period.

Standards for objective 1:

1. *Standard*: Nearly all trainees successfully complete the program. Index: 90 % of trainees complete the program
2. *Standard*: Most trainees embark on careers in biomedical informatics research or continue their education. Index: 75 % or more continue in a career or obtain additional graduate training.
3. *Standard*: The majority of graduates of NLM's university-based programs are still pursuing research careers five years after completing their training. Index: 50 % or more are in a research job five years after completing training
4. *Standard*: All trainees have mentors. Index: 100 % have mentors

   (a) Who actively conduct research in informatics. Index: receive grant during the award period.
   (b) Who have experience as a mentor.
   (c) Who engage the trainee in substantial research projects. Index: co-author of papers.

After establishing the draft evaluation framework, data were extracted from 17 grant applications received in 2001 and 2006, along with available progress reports during that period. Analysis demonstrated that NLM's training programs exceeded the proposed benchmarks in many areas, and that data were not available for some candidate benchmarks. Highlights of the findings for Objective 1:

- 93.5 % of NLM's trainees supported between 1991 and 2005 completed their training ($n=693$)
- 76 % were still in the field in academic, industry or healthcare positions, based on position titles. Former postdoctoral trainees were more likely to be faculty members (44 % compared to 31 % for predoctoral trainees). About 15 % were pursuing additional training. Others were in government or other agencies.
- The majority of trainees (82 % in 2001, 61 % in 2006) had mentors who were principal investigators of active research grants, and 74.9 % of those who had published had published with their mentor.

After going through each objective and the findings, the evaluation framework's standards and index measures were adjusted. In some cases, a standard was restated due to the unavailability of reliable index data. Any recipient of an NIH training grant can be assured that some similar evaluation framework is being used by the funding agency both at the individual award level and the programmatic level.

## Summary

Training for careers in biomedical informatics has entered its fourth decade at NLM. From its early roots in healthcare information management, the field of biomedical informatics evolved and grew into an academic discipline with many distinct subdomains. Today, biomedical informatics trainees include physicians, biologists, nurses, public health administrators, librarians, computer scientists and many others. Graduates move into careers in academic centers, industry, government, small business and other public and private agencies. NLM's experience with its programs and their graduates suggests that the most successful informatics training programs have these qualities:

- Offer courses and experiences that address the broad array of skills and knowledge that can apply to different types of biomedical informatics careers
- Update course content and teaching methods regularly, including use of teaching technologies and self-guided learning as appropriate
- Require didactic, quantitative and computational elements, plus practicum experience for every trainee
- Provide easy access to training in management skills such as budgeting, grant writing or managing a research team
- Have core faculty who employ hands-on mentoring that involves face-to-face meetings, annual written evaluations and career counseling. Provide dual mentors when feasible or needed for the trainee's research
- Form collaborations that provide synergy for program goals, with collaborating faculty in relevant academic departments, business or government who can teach and mentor informatics trainees
- Offer 'identity-building' group experiences for their trainees, such as joint retreats, journal clubs and required participation in speaker series. This is especially important if trainees are located in several academic departments or if they are in distinct tracks within a department.
- Require trainees to make regional or national presentations and write peer-reviewed publications during their training period.
- Gather data in an ongoing way that are useful for evaluation and tracking of individual progress, of curriculum quality, faculty strength.
- Arrange for external evaluation of the academic program on a regular basis, at least every five years

**Key Take-Away Points**

Features of strong Informatics programs:

- Interdisciplinary content, with coverage of information science, cognitive science and knowledge of one or more domains of biomedicine
- A core curriculum of required courses emphasizing informatics concepts and methods and state-of-the-art technology assessments

- Electives providing opportunities for advanced training in informatics fields
- Individual research experience for each trainee, including assistance for trainees in selecting appropriate research projects
- Exposure to the informatics of subdomains such as public health or bioinformatics
- Effective programs for recruiting and retaining a diverse pool of trainees
- Approaches for evaluating program success

## References

1. Cummings MM, Corning ME. The Medical Library Assistance Act: an analysis of the NLM extramural programs, 1965–1970. Bull Med Libr Assoc. 1971;59(3):375–91.
2. National Library of Medicine 135th anniversary report, 1836–1971. Bethesda: U.S. Department of Health, Education, and Welfare. Public Health Service, National Institutes of Health; 1972. DHEW Publication No. (NIH) 72–256.
3. Miles WD. A history of the National Library of Medicine. The nation's treasury of medical knowledge. Bethesda: U.S. Department of Health, Education, and Welfare. Public Health Service, National Institutes of Health, National Library of Medicine; 1982. NIH Publication No. 85–1904.
4. National Library of Medicine. Programs and Services. Fiscal Year 1978. Bethesda: U.S. Department of Health, Education, and Welfare. Public Health Service, National Institutes of Health. National Library of Medicine; 1979. DHEW Publication No. [NIH] 79–256.
5. National Library of Medicine. Programs and Services. Fiscal Year 1983. Bethesda: U.S. Department of Health, Education, and Welfare. Public Health Service, National Institutes of Health. National Library of Medicine; 1984.
6. National Library of Medicine. Programs and services. Fiscal Year 1984. Bethesda: U.S. Department of Health, Education, and Welfare. Public Health Service, National Institutes of Health. National Library of Medicine; 1985.
7. National Library of Medicine. NLM's university-based biomedical informatics research training programs. 2012. http://www.nlm.nih.gov/ep/GrantTrainInstitute.html. Accessed 18 Feb 2013.
8. National Library of Medicine. National Library of Medicine: Institutional Training Grants for Informatics Research RFA-LM-01-001. 2001.
9. National Library of Medicine. Institutional Grants for Research Training in Biomedical Informatics (T15) RFA-LM-06-001. 2006.
10. National Library of Medicine. NLM Institutional Training Grants for Research Training in Biomedical Informatics (T-15) RFA-LM-11-001. 2011.
11. National Library of Medicine. Medical Informatics Research Training Programs RFA-LM-96-002. 1996.
12. The Biomedical Information Science and Technology Initiative. Working Group on Biomedical Computing, Advisory Committee to the Director. Bethesda: National Institutes of Health. 3 June 1999. Online publication accessed 7/6/2013 at http://acd.od.nih.gov/agendas/060399_Biomed_Computing_WG_RPT.htm
13. National Library of Medicine. National Library of Medicine Programs and Services Fiscal Year 2001. Bethesda: Department of Health and Human Services Public Health Service; 2001.
14. National Library of Medicine. National Institutes of Health National Library of Medicine Programs and Services FY 2005. Bethesda: Department of Health & Human Services; 2005.

15. National Library of Medicine. NLM fellowship in applied informatics PA-92-90. 1992.
16. Biomedical Research Workforce Working Group Report. A Working Group of the Advisory Committee to the Director. Bethesda: National Institutes of Health, 14 June 2012. Online publication accessed 7/6/2013 at http://acd.od.nih.gov/Biomedical_research_wgreport.pdf
17. National Institutes of Health – Office of Extramural Research. Funding opportunities and notices. http://grants.nih.gov/grants/guide/index.html (2013). Accessed 18 Feb 2013.
18. National Institutes of Health. NIH Home – Institutes, Centers and Offices. 2013. http://www.nih.gov/icd/. Accessed 19 Feb 2013.
19. Humanitas Inc. Components of a model program. Analysis and evaluation of NLM's Informatics Research University Training Programs. 2010. Phase 3, Deliverable 5. 3 Dec 2010.
20. Humanitas Inc. Analysis and evaluation of NLM's Informatics Research University Training Programs. 2008. Deliverable 7, Phase 1 final report.
21. Rice University. Professional Development for Graduate Scientists – a training materials repository. 2012. http://cnx.org/content/m37008/latest/.
22. National Institutes of Health. Advisory Committee to the Director – Working Group on data and informatics. 2011. http://acd.od.nih.gov/DIWG.htm. Accessed 19 Feb 2013.
23. Kulikowski CA, Shortliffe EH, Currie LM, Elkin PL, Hunter LE, Johnson TR, Kalet IJ, Lenert LA, Musen MA, Ozbolt JG, Smith JW, Tarczy-Hornoch PZ, Williamson JJ. AMIA Board white paper: definition of biomedical informatics and specification of core competencies for graduate education in the discipline. J Am Med Inform Assoc. 2012;19(6):931–8. doi:10.1136/amiajnl-2012-001053.
24. Draft Report of the Advisory Committee to the Director Working Group on Diversity in the Biomedical Research Workforce. Bethesda. 13 June 2012. Online publication accessed 7/6/2013 at http://acd.od.nih.gov/Diversity%20in%20the%20Biomedical%20Research%20Workforce%20Report.pdf
25. "About ES CareerTrac." A web-based explanatory overview of a trainee tracking system developed and used by the National Institute of Environmental and Health Sciences and the Fogarty International Center. National Institutes of Health. Online publication accessed 7/6/2013 at https://careertrac.niehs.nih.gov/about.gsp