

Don Lindberg's influence on future generations: The U.S. National Library of Medicine's biomedical informatics research training programs

Robert A. Greenes^{a,*}, Valerie Florance^b and Randolph A. Miller^c

^aArizona State University, Phoenix, AZ, USA

^bU.S. National Library of Medicine, Bethesda, MD, USA

^cVanderbilt University School of Medicine, Nashville, TN, USA

Abstract. Through his visionary leadership as Director of the U.S. National Library of Medicine (NLM), Donald A. B. Lindberg M.D. influenced future generations of informatics professionals and the field of biomedical informatics itself. This chapter describes Dr. Lindberg's role in sponsoring and shaping the NLM's Institutional T15 training programs.

Keywords: Donald A.B. Lindberg, U.S. National Library of Medicine, biomedical informatics training

In memory of Milton Corn M.D. 1928–2021. Director of the Extramural Programs Division, U.S. National Library of Medicine 1990–2009; Deputy Director for Research and Education 2009–2021.

1. Introduction

Donald A.B. Lindberg M.D. influenced successive generations of trainees in the field of biomedical informatics during three decades as Director of the U.S. National Library of Medicine (NLM). This chapter describes how Dr. Lindberg's visionary leadership both enabled and shaped training in biomedical informatics from 1984–2015. The chapter will not describe in detail the NLM Informatics Short Course or separate training opportunities offered by NLM to librarians or by the National Center for Biotechnology Information (NCBI) [1].

Prior to Dr. Lindberg's arrival at NLM, the Medical Library Assistance Act of 1965 had authorized NLM to provide grants for post-baccalaureate academic training in medical library sciences, in related fields pertaining to sciences related to health, and in the field of the communication of information [2]. By 1969, an annual budget of \$1.3 million was allocated among 11 universities for development of graduate and postgraduate training programs for health science information and communication specialists [2]. Four of those programs supported 11 Ph.D. candidates for careers in health information research.

*Corresponding author: Robert A. Greenes, Emeritus Professor, Arizona State University, Phoenix, AZ, USA. E-mail: greenes@asu.edu.

In 1972, NLM and the Bureau of Health Manpower Education cooperated to support the training of physicians and other health scientists in the use of computer technology for medical education and the provision of health care [2]. Those early training programs differed significantly from current NLM T15 grant-based training programs, as the 1976 NLM Annual Report clearly explained:

“The National Library of Medicine supports the training of health scientists in computer technology. The ultimate goal ... is to promote the complete and effective integration of computer technology into all phases of clinical medicine, teaching, practice, and research. Except for a few unusual people, most do not find it possible to master two disciplines well enough to provide innovative leadership in either ... Therefore, the most practical projects to be supported are those for teachers or potential faculty members in the health sciences. By incorporating new ideas into their teaching, they will multiply the process of dissemination. Depending on the academic background and experience of the faculty member being trained, one or two years of guided study should suffice to provide him with insight into the potential of computer techniques. The training should equip him to recognize where computers can be of assistance in the solution of medical problems, to communicate his needs effectively to computer consultants, and to understand and use the computer assistance provided [3]”.

By 1980, grants for 76 individual trainee positions (36 pre-doctoral, 39 post-doctoral) were awarded via this program to 10 institutions [2]. In 1982–83, the NLM Board of Regents decided to sponsor institutional grants that focused on research training rather than the previous mechanism of grants that supported training of specific individuals [4]. In 1984, five awards were made in the new Health Computer Sciences Research Training program (detailed below) [5].

When Dr. Lindberg was sworn in as NLM Director in October 1984, he had already gained a significant understanding of NLM's programs through earlier service on the NLM Board of Scientific Counselors (1983–84) and the NLM Biomedical Library Review Committee (1976–80). At Dr. Lindberg's first NLM Board of Regents (BOR) meeting in October 1984, his presentation stimulated a discussion among BOR members regarding the need for long-range planning [6]. The BOR then formally requested creation of such a plan. Central to Don's emphasis on training was the NLM Long-Range Plan of 1986. The presentation of the plan to the BOR in 1986 stated that: “NLM should assume leadership in the medical Informatics area, become a ‘National Institute for Medical Informatics;’ ... training in medical Informatics should be encouraged in schools of the health professions [7]”.

2. Don Lindberg's ongoing personal influence on NLM Institutional Training Programs

Dr. Lindberg saw the NLM Institutional Training Program as a commitment and opportunity to build the nascent community of informatics professionals. NLM sponsored an annual training meeting that brought faculty and trainees from all the NLM-funded training sites together. The NLM annual training meetings were designed to encourage faculty and trainees to interact socially, to exchange ideas, and to build a sense of community among them. Dr. Lindberg also viewed the annual training meetings as a forum to introduce faculty and trainees to the opportunities provided by NLM for support of their future research.

The formal components of the earliest NLM annual training meetings included a gathering of all attendees in an auditorium at NLM with prominent plenary session speakers, including Don himself as the leadoff presenter. Subsequent days of the training meeting were reserved for lecture and poster presentations by trainees. Don also hosted a break-out session for the Training Program Directors at each

meeting to review operational issues at NLM and at each site, to present future funding outlooks, and to exchange suggestions for changes/improvements.

In his keynote presentations, Don would summarize the state of the field, highlight current initiatives at NLM, and point out what he saw as future research prospects. For example, during the 1990 Annual NLM Training Meeting plenary session, Don exhorted everyone present (faculty and especially trainees) to consider expanding their individual careers and their training programs to encompass bioinformatics. That was the first time many individuals in the audience heard of NLM's recent commitment to the domain - embodied in the new National Center for Biotechnology Information [8].

An important social component of the meetings was the annual picnic, typically hosted by Don and his wife, Mary. In the earliest years, they hosted the afternoon and evening get together with a cookout at their Potomac area home. Many of the trainees enjoyed swimming in the Lindberg's back-yard pool. When the program expanded over the years to outgrow the capacity of the Lindberg's house, the venue for the social event changed to the picnic grounds behind the (then) Bethesda Naval Hospital at the Uniformed Services University of the Health Sciences Medical Center. Volleyball games and horseshoe tossing contests included program directors, faculty, and fellows. By combining both formal meetings and social events, these became very meaningful events to the trainees. In addition, trainees were thrilled to meet and interact informally with Don and with training directors and faculty from other universities.

After the initial five years of annual training meetings had been hosted at NLM, a decision was made to alternate the meetings between NLM and one of the training program sites that volunteered to serve as host for a given year. The host sites spanned the U.S., giving trainees the opportunity to see firsthand how informatics was manifested in different environments. Each host site also took pride in sponsoring a picnic or dinner at a noteworthy local attraction. For example, when the Harvard-MIT program hosted the meeting, the social event included a clambake in the courtyard outside the Harvard medical library. When Yale hosted the meeting, dinner attendees were treated to a Yale Glee Club performance.

The initial training programs focused on clinical informatics. It was Don's foresight that contributed greatly to broadening the field eventually into biomedical informatics, by calling attention to the growing work in molecular science and genomics and activity centered around bioinformatics at NIH. During Don's tenure at NLM the scope of the programs expanded to include high-performance computing and computational complexity, vocabulary and terminology, public health, and toward the end of his time at NLM, biomedical data science.

3. Growth, diversification, and evolution of the NLM institutional training programs

One of the features of the NLM-sponsored training programs from 1984 onward was the specific focus on research training [2]. Trainees were expected to undertake academic study at the doctoral or postdoctoral level, and to prepare for careers that would advance the field. This was very important for a growing field in which prior leaders had been largely self-taught through ad hoc coursework and on-the-job experience. At the time of the initial training programs, it was already recognized that biomedical informatics was a cross-disciplinary field that drew on several other disciplines, including computing, psychology, statistics, biomedical science, and clinical medicine, among others. Curricula were just beginning to be developed, as well as formal degree programs. Some historical Training Program highlights (derived from the NLM Annual Reports and Minutes of the NLM Board of Regents Meetings) follow below.

Prior to Don Lindberg's arrival at NLM, the first institutional research-focused grants made through the new Health Computer Sciences Research Training program supported five sites, as noted in the introduction. The 1984 awards included: University of California, San Francisco School of Medicine (PI: Marsden S. Blois M.D., Ph.D.); University of Minnesota School of Medicine (PI: Lael Gatewood Ph.D.); Harvard University (PI: Robert A. Greenes M.D., Ph.D.); Tufts-New England Medical Center (PI: Stephen G. Pauker M.D.) and Stanford University School of Medicine (PI: Edward H. Shortliffe M.D., Ph.D.) [5]. The five institutions enrolled 21 individuals initially, subsequently ramping up to more than 50 trainees per year. The awards primarily focused on postdoctoral research training, including individuals with prior M.D.s, some with Ph.D.s, as well as individuals in M.D.-Ph.D. programs. In later years, programs could request funds for three-month short-term trainees, such as medical students, to give them exposure biomedical informatics as a possible career path.

NLM's training programs were re-competed in 1987, to have a broader biomedical informatics research focus, to include biotechnology informatics as well as clinical informatics [9]. The program was expanded to eight institutions, with support for more than 70 trainees, still predominantly postdoctoral in emphasis, but now providing support for doctoral candidates as well. A re-competition in 1992 resulted in 10 training sites, some of which were now multi-institutional. Trainee count expanded to more than 100 trainees per year. The National Cancer Institute (NCI) contributed support for four trainees per year [10]. The 1997 round of funding resulted in awards to 10 training sites again, with several again multi-institutional [11]. Trainee count was now over 150 per year. Some of the NLM training slots focused on radiation oncology funded by NCI and for dental informatics funded by the National Institute of Dental Research (NIDR). The final year of NCI support was 1999.

In 2002, 18 programs were funded for next five-year interval. More programs emphasized bioinformatics, and with partial support from the National Institute of Biomedical Imaging and Bioengineering (NIBIB), one program emphasized imaging informatics. More than 270 trainees were being supported annually by the end of this award period. In 2005, the Robert Wood Johnson Foundation provided support to increase training in public health informatics, resulting in supplemental support to four existing training sites to develop formal training tracks in public health informatics and for support of trainees in these tracks. In 2007, the program continued to support more than 250 trainees per year.

Collectively, the programs emphasized training in health care informatics (14 programs), bioinformatics and computational biology (14 programs), clinical research translational informatics (13 programs), and public health informatics (10 programs). Co-funding was received from the National Institute of Dental and Craniofacial Research, which supported training in dental informatics, and from the Robert Wood Johnson Foundation (RWJ) for training in public health informatics.

NLM's funding of its grant programs was downsized in 2011 [12]. This limited the number of training programs funded in the re-competition of 2012 to a maximum of 15 programs with 15 trainees or less at each site. This also included a shift to include more emphasis on bioinformatics and translational bioinformatics. The latter domains were expected to highlight connections to human health. In addition to the one new and 14 continuing sites selected, six existing programs that were not awarded new five-year grants received additional funding for one year for doctoral students early in their studies. Collectively, the continuing programs emphasized training in: health care informatics (12 programs); bioinformatics and computational biology (12 programs); clinical research translational informatics (12 programs); and public health informatics (10 programs). The National Institute of Dental and Craniofacial Research expanded support for trainees in dental informatics from two to nine trainees.

The competition that led to the 2017 training awards was influenced by NLM's participation in several NIH Big Data to Knowledge (BD2K) informatics and data science training initiatives [13].

The National Institute of Environmental Health Sciences (NIEHS) provided funding for environmental informatics trainees. The 14 university-based training programs supported 117 pre-doctoral trainees and 93 postdoctoral trainees.

As of 2018, the actively supported NLM training sites and their years of continuous support included: University of Pittsburgh (32); Yale University (32); Stanford University (35); University of Utah (22); University of Wisconsin-Madison (17); Indiana University Purdue University at Indianapolis (2); University of Buffalo (2); Vanderbilt University (17); Oregon Health & Science University (32); Columbia University (27); University of Colorado (12); Rice University (27); Harvard Medical School (27); University of North Carolina (2); University of California, San Diego (7); and University of Washington (17). In the interim from 1984 to 2018, the following sites (not listed above as initial or current sites) hosted NLM-sponsored biomedical informatics training programs for varying durations: Washington University, St. Louis; Duke University; Indiana University/Regenstrief Institute; University of California, Irvine; University of California, Los Angeles; Johns Hopkins University; University of Missouri, Columbia; University of South Carolina; MIT; and Baylor University [13].

The Training Grants were funded through the Medical Library Assistance Act portion of NLM's budget. NLM's training programs were unique at NIH in providing up to five years of funding for a predoctoral trainee, and supplemental funds for travel, training-related expenses, and health insurance. In an early period, NLM even supplemented the stipends for trainees with computer science training to lure them into informatics. Of note, NIH rules and previous legislation had excluded NLM from using the Kirschstein T32 training grant mechanism. In years when the NLM experienced funding constraints,

Dr. Lindberg did his utmost to preserve funding to the training programs. An observation about the evolution of the programs is that, although its size and scope was dependent on NLM's funding, the allocation of funds and preservation of those funds among competing priorities owe much to Don Lindberg's commitment to the goals of training the next generation. Don also approved the various partnership agreements that resulted in co-funding of training positions by other agencies over the years. The continual reshaping and expansion of focus, plus the ongoing commitment to research training are testimony to Don's leadership and vision, and his extramural grants program staff. They recognized important new areas that expanded the scope of biomedical informatics, often before many of their colleagues were fully aware of them.

4. Outcomes of the NLM training programs

Over the years, as more biomedical informatics degree programs were offered at various universities in the U.S., NLM added a requirement that training-grant sponsored research fellows without a prior doctorate must obtain a doctoral degree, and that postdoctoral fellows must obtain a Master's degree. This served to emphasize NLM's ongoing commitment to advancing the academic, not just the applied, nature of the field. It is probably fair to say that most of the faculty of academically based biomedical informatics programs have come out of such training programs. When looking at the legacy of the training programs in terms of positions of leadership in healthcare organizations, informatics-focused public agencies, and business entities, the graduates of such programs also are well-represented. For example, of the institutional biomedical informatics training programs who received new five-year funding in 2017, one-third of the Principal Investigators/Training Directors were themselves graduates of NLM-sponsored training programs.

NLM sponsored a study of previous trainees' careers during the decade from 1996–2005 [2]. Forty percent had careers in academia; 18 percent had careers in business/industry; 17 percent were employed by healthcare organizations; 10 percent chose to undertake further training; and one percent were government employees. Between 1996 and 2005, 507 NLM trainees completed their training (37 percent were female); 67 percent of them had worked on research with funded researchers, and 38 percent of them published at least once with their mentor before they left the program. On average, each program received 54 applications and made offers to 24 percent of applicants; there was a 90 percent retention rate; 81 percent of core faculty had active research grants, and 38 percent co-authored papers with trainees.

Between 1996 and 2006, 51 percent of NLM trainees had been lead or co-author on one or more articles, conference papers, abstracts, or books, and they had produced 1,452 peer-reviewed publications.

NLM identified the following as valuable components of its training programs (and associated training meetings): community building (training the next generation of informaticians); informal exchange of ideas at annual meetings; fostering collegiality among trainees at annual meetings; trainees visiting NLM as a future sponsor for their research; introduction of trainees to NLM personnel; enhancement of multi-axial diversity in the field; and scientific and educational exchange of ideas (trainee presentations at annual meetings, training directors meeting, enhancement of understanding of diverse disciplines). The benefits to institutions from NLM's training programs included: formation and growth of academic units; enhanced breadth of coverage within faculty units; and talented graduate students advancing faculty's research.

In May 2012, Extramural Programs (EP) Director Dr. Valerie Florance reported to the NLM Board of Regents that NLM's training program consumed roughly 25–30 percent of the total NLM EP budget, in comparison to the average amount that NIH overall spends on research training, which was then less than five percent [14].

5. Conclusion

The NLM Institutional Research Training Programs have been key to the expansion of biomedical informatics both academically and in organizations that rely on individuals well trained in the broad aspects of the field and its scientific underpinnings. As the discipline matured over the past four decades, NLM's training programs mirrored ongoing developments by expanding their scopes to include emerging foci such as bioinformatics, public health informatics, imaging informatics, research computing, data sciences, and other advances. During this period, NLM-supported informatics training programs also increased the academic credentials required for trainees to complete. In more recent years, this included masters and doctoral degrees to prepare trainees for research careers.

The significant allocation of funding to training and the expansion of scope over the years owe much to Don Lindberg's vision and leadership. He continued to challenge the programs with new opportunities and NLM provided the funds to respond to those opportunities. Dr. Lindberg imbued the program with a strong sense of community, which has persisted and grown through the present time.

References

- [1] J.J. Cimino, The biomedical informatics short course at Woods Hole/Georgia: training to support institutional change. in: *Transforming Biomedical Informatics and Health Information Access: Don Lindberg and the U.S. National Library of Medicine*, B.L. Humphreys, R.A. Logan, R.A. Miller and E.R. Siegel (eds), IOS Press, Amsterdam, 2021.
- [2] V. Florance, R.A. Miller and M. Corn, NLM training grants. in: *Panel Presentation, National Library of Medicine at its 175th Anniversary: What it has Accomplished*. AMIA Fall Symposium, Washington, DC, 2010.

- [3] National Library of Medicine (U.S.). Programs and services fiscal year 1976, p 42. [cited 2021 August 19]. Available from: <https://www.nlm.nih.gov/hmd/manuscripts/nlmarchives/annualreport/1976.pdf>.
- [4] National Library of Medicine (U.S.). Programs and services fiscal year 1983 [cited 2021 June 21]. Available from: <https://www.nlm.nih.gov/hmd/manuscripts/nlmarchives/annualreport/1983.pdf>.
- [5] National Library of Medicine (U.S.). Programs and services fiscal year 1984 [cited 2021 June 21]. Available from: <https://www.nlm.nih.gov/hmd/manuscripts/nlmarchives/annualreport/1984.pdf>.
- [6] National Library of Medicine (U.S.). Board of Regents. Minutes of Board of Regents Meeting. [Internet] October 1984 [cited 2021 June 21]. Available from: <https://www.nlm.nih.gov/hmd/manuscripts/nlmarchives/bor/1984.pdf>.
- [7] National Library of Medicine (U.S.). Board of Regents. Minutes of Board of Regents Meeting. [Internet] June 1986 [cited 2021 June 21]. Available from: <https://www.nlm.nih.gov/hmd/manuscripts/nlmarchives/bor/1986.pdf>.
- [8] D.R. Masys and D.A. Benson, Don Lindberg and the creation of the National Center for Biotechnology Information. in: *Transforming Biomedical Informatics and Health Information Access: Don Lindberg and the U.S. National Library of Medicine*, B.L. Humphreys, R.A. Logan, R.A. Miller and E.R. Siegel (eds), IOS Press, Amsterdam, 2021.
- [9] National Library of Medicine (U.S.). Programs and services fiscal year 1987 [cited 2021 June 21] Available from: <https://www.nlm.nih.gov/hmd/manuscripts/nlmarchives/annualreport/1987.pdf>.
- [10] National Library of Medicine (U.S.). Programs and services fiscal year 1992 [cited 2021 June 21]. Available from: <https://www.nlm.nih.gov/hmd/manuscripts/nlmarchives/annualreport/1992.pdf>.
- [11] National Library of Medicine (U.S.). Programs and services fiscal year 1997 [cited 2021 June 21]. Available from: <https://www.nlm.nih.gov/ocpl/anreports/fy1997.pdf>.
- [12] National Library of Medicine (U.S.). Programs and services fiscal year 2011 [cited 2021 June 21]. Available from: <https://www.nlm.nih.gov/ocpl/anreports/fy2011.pdf>.
- [13] National Library of Medicine (U.S.). Programs and services fiscal year 2017 [cited 2021 June 21]. Available from: <https://www.nlm.nih.gov/ocpl/anreports/fy2017.pdf>.
- [14] National Library of Medicine (U.S.). Board of Regents. Minutes of Board of Regents Meeting. [Internet] May 2012 [cited 2021 August 18]. Available from: https://www.nlm.nih.gov/od/bor/BOR_May_2012_Minutes.pdf.