Donald A.B. Lindberg and the U.S. National Library of Medicine transformed biomedical and health informatics

Randolph A. Miller[a],[*] and Edward H. Shortliffe[b]
[a]*Vanderbilt University School of Medicine (Emeritus), USA*
[b]*Department of Biomedical Informatics, Columbia University in the City of New York, USA*

Abstract. This overview summary of the Informatics Section of the book *Transforming biomedical informatics and health information access: Don Lindberg and the U.S. National Library of Medicine* illustrates how the NLM revolutionized the field of biomedical and health informatics during Lindberg's term as NLM Director. Authors present a before-and-after perspective of what changed, how it changed, and the impact of those changes.

In memory of Harold M. Schoolman M.D. 1924–2009. Deputy Director for Research and Education, U.S. National Library of Medicine, 1970–1999.

1. Introduction

From the early 1980s to 2015, the computing landscape in the United States evolved from expensive, privately accessed mainframes to ubiquitous, inexpensive desktop computers and powerful handheld devices. During that time, computer input mechanisms progressed from punched cards and keyboards to mouse clicks, touch screens, and speech recognition. Person-to-person phone communication morphed from hard-wired rotary dial telephones (and coin-operated telephone booths!) to compact, handheld lightweight multifunction cellular devices. Remote connections to computing resources advanced from slow, fee-for-service commercial connections using acoustic modems to free high-speed internet access. During the tenure of Donald A.B. Lindberg M.D. as its Director from 1984 to 2015, the U.S. National Library of Medicine (NLM) introduced complementary transformations that radically altered and advanced the landscape of biomedical and health informatics in the United States. The chapters of the Informatics Section of the book, *Transforming biomedical informatics and health information access: Don Lindberg and the U.S. National Library of Medicine* document these transformations [1–12].

Lindberg's early years demonstrated his ability to recognize and solve challenging problems. This foreshadowed his later informatics accomplishments. In *A scientific mind embraces medicine: Donald*

---

[*]Corresponding author: Randolph A. Miller. E-mail: ramiller49@icloud.com.

Lindberg's education and early career, Kingsland and Kulikowski describe Lindberg's formative years at Amherst, Columbia University, and at the University of Missouri [1]. Lindberg pioneered development of laboratory information systems. He installed terminals on hospital wards to deliver laboratory results to clinicians, and correlated microbiology culture results with antibiotic sensitivities. Using the College of American Pathologists' *Systematized Nomenclature of Pathology* to cross-index information from multiple sources, Lindberg oversaw a team that developed a computer-based fact bank. He and a colleague engineered a modified 8-track stereo audio cartridge system that could provide dial-up answers to medical questions posed by underserved rural patients and their families. Lindberg also contributed to development of early computer-based diagnostic decision support tools. His creativity and accomplishments led to his selection as Director of the NLM in 1984. At NLM, Lindberg repeatedly applied, on a much larger scale, the schema he developed in Missouri: identify a novel challenging problem, locate and combine disparate resources needed to address the problem, and provide creative ways to access the information.

In *Lindberg, pioneer in biomedical and health informatics: his involvement in creating professional organizations*, van Bemmel, Ball and Shortliffe reveal another aspect of Lindberg's character - his role as a visionary and highly capable leader [2]. They observe that throughout his career, Lindberg surrounded himself with teams of top researchers and superb staff. Early in his career, he held positions on the Boards of the American Association for Medical Systems and Informatics (AAMSI), and the Symposium on Computer Applications in Medical Care (SCAMC). With colleagues, Lindberg played a critical role in founding the American College of Medical Informatics (ACMI). Those activities, combined with his leadership role at the NLM and his contributions as the American representative to the International Medical Informatics Association (IMIA), prepared him to become the logical choice to serve as the founding President of the American Medical Informatics Association (AMIA). Lindberg not only could unite disparate information resources; he could effectively organize people as well.

2. How NLM laid the foundation for professional and institutional growth in biomedical and health informatics

Lindberg's approach to integrating, cross-referencing, and providing new means of access to biomedical information characterized many of the projects that NLM implemented under his visionary leadership.

The national informatics landscape prior to Dr. Lindberg's arrival at NLM consisted primarily of individual investigator-initiated laboratories in isolated academic settings [13]. A small number of institutions had developed one-of-a-kind clinical information systems that were used locally. Academic medical centers had typically placed their computing facilities under the purview of Chief Financial Officers (CFOs) since billing and accounting were critical to their domain. Most CFOs saw little benefit in diverting computer resources to support healthcare research or medical libraries, or to make direct interventions in care delivery. The academic institutions harboring early informatics faculty members had difficulty understanding how to value and promote individuals whose seemingly opaque work did not fit neatly into existing criteria for advancement.

Several trailblazing NLM projects made key contributions that changed this situation. Their nationwide (and often global) impacts facilitated establishment of biomedical informatics as a modern, professional discipline [3–12,14]. Lorenzi and Stead describe the development and evolution of the Integrated Academic Information Management Systems project in *NLM and the IAIMS Initiative: cross-institutional academic/advanced systems contributing to the evolution of networked information and resources* [3]. Following a 1962 study that indicated academic medical libraries were underutilized, the Medical Library

Assistance Act (MLAA) of 1965 authorized NLM to provide resources to medical libraries and to create a network of regional medical libraries. The 1982 Cooper–Matheson report encouraged further action; NLM issued IAIMS contracts to four academic medical centers just before Lindberg's arrival. Under Lindberg and Richard T. West M.L.S., the NLM IAIMS Program Officer, IAIMS initiatives expanded and evolved.

The IAIMS calls for applications encouraged institutions to comprehensively and strategically plan how to interconnect people and information in clinical facilities, the medical library, research laboratories, and the educational arena. The goal was to address critical communication and information needs. The integration that IAIMS fostered was both physical (via networks) and intellectual (via computer applications providing new functionality). The IAIMS projects taught administrators, teachers, researchers, and students that informatics as a discipline had the potential to augment and improve their work. The lack of appreciation for informaticians on academic faculties dissipated, and institutions began to seek additional faculty trained in the field. The medical centers that received IAIMS grants during Dr. Lindberg's tenure at NLM became national leaders in informatics. They were widely emulated at other sites. Thus, the NLM IAIMS projects advanced academic institution's informatics infrastructures in profound ways, far beyond the set of IAIMS-funded sites.

Before the NLM began to issue research-oriented T15 training grants in 1984, a small number of institutions had begun to offer one-of-a-kind approaches to training individuals interested in informatics careers - via ad-hoc interdisciplinary degrees or custom apprenticeships. Greenes, Florance, and Miller describe how NLM established a uniform foundation for ongoing growth in the field in *Don Lindberg's influence on future generations: the U.S. National Library of Medicine's biomedical informatics research training programs* [4]. By 2017, shortly after Don Lindberg's retirement as NLM Director, 24 universities/institutions had received T15 training grants, and more than 1000 future members of the field had received NLM-sponsored training. Many of them later became leaders in the field at local and national levels.

Moreover, the NLM training grants promoted crucial interactions among new entrants into the field. The annual grant-related training meetings established a camaraderie among trainees and introduced them to senior faculty leaders from other institutions. The meetings acquainted trainees with NLM staff and showcased NLM programs as potential means of future support. The training meetings increased awareness of the national scope of informatics activities.

By the early 1990s, interest in biomedical informatics careers had grown substantially, and extended far beyond sites having IAIMS or T15 training grants. Yet no national-level mechanisms existed to introduce uninitiated administrators, faculty, and librarians to the field. Lindberg strongly believed in using outreach programs to promote the adoption of NLM's initiatives. As James J. Cimino M.D. describes in *The biomedical informatics short course at Woods Hole/Georgia: training to support institutional change*, NLM developed a series of week-long informatics short courses [5]. These intensive courses were offered at Woods Hole, Massachusetts and, subsequently, at Augusta University in Georgia with a goal of training potential informatics change agents. Participants had to apply to be selected to attend. NLM paid for attendees' tuition, room, and board. Like the T15 training grant meetings, the short courses facilitated long-lasting collaborations among participants. Subsequently, many colleges and universities developed informatics course offerings patterned after the NLM short courses — including AMIA's 10 × 10 courses.

Before Lindberg became Director of the NLM, an individual entering the new academic discipline of biomedical informatics had few foreseeable options for obtaining ongoing long-term financial support. In *NLM's sponsorship of research in biomedical informatics (1985–2016)*, Kuo and Ohno–Machado review NLM's extramural R01 grant sponsorship during Lindberg's tenure [6]. Over three decades, NLM

research grants evolved to support the field as it transitioned from a clinical informatics focus to add support for research in translational bioinformatics. The NLM Extramural Programs' annual budget kept pace with the growing number of individuals entering the field. It grew from $7.5 million in 1983 to $42 million in 2015. In addition, following NLM's lead, other NIH institutes began to sponsor informatics research and development projects.

Collectively, NLM's institutional IAIMS programs, T15 training grants, informatics short courses, and extramural research support created a vital infrastructure that supported the future growth of the field.

3. Examples of how the new NLM information infrastructure enhanced research and clinical care

During Lindberg's tenure, the NLM developed critical international information resources and novel informatics methods to access them. As Mo and Denny indicate in *The U.S. National Library of Medicine's impact on precision and genomic medicine*, Dr. Lindberg realized that the key to transforming biomedical research and clinical practice was creation of cross-indexed, interconnected databases with advanced access tools [7]. The tools were made publicly available at no cost to users.

In 2021, clinicians access NLM bibliographic, toxicology, public health, and genomic databases (among others) to diagnose and treat patients' disorders. Advanced cancer centers use NLM databases to develop and implement genetically targeted cancer therapies. The NLM's resources played key roles in addressing the COVID-19 epidemic of 2019–2021. Such capabilities were barely imaginable in 1984. As a result, the NLM fundamentally changed the scope of available problem-solving approaches for biomedical researchers and clinicians and created new resources that could be mustered to address evolving problems. These enabled advances in genomic and personalized medicine.

McDonald and Humphreys, in *The U.S. National Library of Medicine and standards for electronic health records: one thing led to another*, explain how NLM evolved from an early position of eschewing ownership of non-bibliographic standards, to becoming the major U.S. Government maintainer and disseminator of healthcare-related standards [8]. These include resources developed during the UMLS project (see below) and others: LOINC, SNOMED, RxNorm, Structured Product Labels, and the DailyMed drug information distribution system. McDonald and Humphreys illustrate how the NLM's standards efforts enabled development of health information exchanges (HIEs), among many other important applications. An HIE creates a virtual patient record system - enabling clinicians at one of many sites where a patient has been seen to view an integrated composite patient record drawing data from all sites. It is only by mapping disparate data at each participating HIE site to NLM-sponsored standards that a monolithic, integrated patient view becomes possible.

4. How NLM created an information infrastructure that enhanced progress in the field

Four key NLM projects provided the foundation for the above-mentioned advances in clinical care and research. Humphreys and Tuttle describe the first of these in *Something new and different: the Unified Medical Language System* [9]. Masys and Benson in *Don Lindberg and the creation of the National Center for Biotechnology Information* document the second critical development (NCBI), which has had world-wide impact in promoting "omics" research [10]. The third development, the High-Performance Computing and Communications (HPCC) project, is described by Ackerman, Howe and Masys in *Don Lindberg, high performance computing and communications, and telemedicine* [11]. Finally, Ackerman,

in *The Visible Human project* [12], details development of a resource that had implications well beyond anatomy representation.

The UMLS project began by enlisting university-based academic informatics teams into one of the earliest large-scale collaborations in the field. Key participants on the highly talented NLM project team included Lindberg, Harold M. Schoolman M.D., Betsy Humphreys M.L.S., Lawrence C. Kingsland Ph.D., Peri L. Schuyler M.L.S., Alexa T. McCray Ph.D., Daniel R. Masys M.D., and William T. Hole M.D. An initial round of individual demonstration projects gave way to coordinated efforts to build and maintain the main component of the project, the UMLS Metathesaurus. Lindberg envisioned it as enabling "the retrieval and integration of information from disparate electronic sources, e.g., patient records, biomedical literature, knowledge bases [9]".

Despite the talents of the assembled UMLS team, Lindberg realized that the NLM did not have the expertise or personnel required to build and maintain a single source "official" vocabulary for biomedicine which could support all aspects of clinical practice and basic biomedical research. Instead, NLM adopted the approach of interconnecting frequently-used, well-organized and well-maintained vocabulary systems from both internal NLM sources (e.g., Medical Subject Headings - MeSH, and TOXLINE) and external sources (e.g., the International Classifications of Diseases, SNOMED, Current Procedural and Terminology, the Diagnostic and Statistical Manual of Mental Disorders, etc.).

The NLM engaged the services of Lexical Technology Inc. to carry out advanced processing of the raw component databases. The goal was to identify unique intellectual concepts no matter how the concepts' names varied across each UMLS source vocabulary. The Metathesaurus then designated a canonical name for each concept, and assigned a unique identifier that would not change even if the canonical name evolved. In turn, the project linked each canonical concept term to synonyms and lexical variants and maintained mappings back to the source vocabularies. The NLM subsequently "decided to create a separate UMLS Semantic Network, consisting of high-level Semantic Types or categories, e.g., Medical Device, Anatomic Abnormality, and the sensible relationships among them [9]".

The original Methathesaurus content included 64,000 concepts and 200,000 concept name variants. Methathesaurus content expanded 70-fold in size from 1990 to 2021. Applied UMLS projects world-wide number in the hundreds to thousands and range from supporting electronic health record systems to natural language processing applications, cataloging contents of medical school curricula and patient safety reports, linking patient charts to bibliographic references, and extracting health-related information from social media - among many others. Mo and Denny emphasize the significant impact of the UMLS in today's clinical practice environment and in biomedical research [7].

The NCBI grew out of a 1986 NLM Long-Range Plan visioning session whose charge was to improve the retrieval of factual information from databases. Three current and future Nobel laureates were in attendance. They indicated to Lindberg that current methods of accessing relevant databases in molecular biology and genetics research amounted to a modern Tower of Babel. Each database had a unique organization, a unique terminology, and unique access methods.

"The incompatibility of these closely related scientific resources thwarted a researcher's ability to use similarities and insights from one database to explain findings recorded in another and contrasted with the scientific literature where a single experiment might produce data that was then included in several disparate databases [10]". The 1987 NLM Long Range Plan included the recommendation: "Immediately establish an intramural and extramural program for biotechnology information. The intramural component should be a National Center for Biotechnology Information, to serve both as a repository and distribution center for the growing body of knowledge and as a laboratory for developing new information analysis and communications tools essential to continued advancement in this field [10,15]".

Lindberg assembled a team that included Dan Masys M.D., Dennis Benson Ph.D., and NLM Deputy Director Kent Smith. Their preparatory work culminated with legislation authorizing creation of the NCBI in November, 1988. David Lipman M.D. was hired as NCBI Director in 1989. The legislative charge given to NCBI included: develop automated systems to store, retrieve, and analyze information related to molecular biology, biochemistry, and genetics; determine how to represent and manipulate knowledge about biologically important molecules; enable biotechnology researchers to access the aforementioned systems and information; and coordinate collection of biotechnology information internationally [10]. The rest, as the saying goes, is history.

"For 28 years, until retiring in 2017, Lipman and his fellow NCBI leaders translated NLM's interest in advancing molecular science into tangible and widely used resources and tools for researchers worldwide. By all measures, the organization has exceeded the goals originally envisioned by Don Lindberg and the Long-Range planners, and its services have become woven into the fabric of 21st century science, continuing to catalyze biomedical research on a global scale [10]". The PubMed interface developed by NCBI integrates and cross-references, among other things, access to NLM's bibliographic databases, OMIM, GenBank, dbSNP, and ClinVar, and MedlinePlus consumer-oriented health information. MedlinePlus' components include *Genetics* reference materials, general *Health Topics*, information on *Drugs & Supplements*, and information about *Medical Tests*. Again, Lindberg's schema of integrating, cross-linking, and providing new ways to access biomedical information paid significant dividends.

The HPCC project followed the theme of uniting resources on a much grander national scale. Ackerman, Howe, and Masys note that: "U.S. Government support for advanced computing in the 1950s and 1960s slackened in the 1970s [11]". In contrast, the High Performance Computing Act, sponsored by Senator Al Gore, became law in 1986, with the goal of learning how to enhance computer network speeds, increase connectivity on a national level, and improve access to supercomputing centers. In 1991, the President's Office of Science and Technology Policy launched the HPCC program, and Don Lindberg was named Director of its National Coordination Office, concurrent with his ongoing position as NLM Director. Among its health-related projects, HPCC advanced "biomedical imaging, molecular biology, and molecular dynamics, and included NLM's Visible Human Project … and digital libraries [11]".

Ackerman tells the story of the Visible Human Project from his perspective as its prime mover [12]. His early focus at NLM was microcomputer-based education. After discussions with faculty members at the University of Washington about the shortcomings that medical students experience during cadaver dissections, Ackerman posed the question, "Would focusing on anatomy provide the example that would get medical schools to adopt interactive technologies? [12]".

Dan Masys, M.D., as then-Director of NLM's Lister Hill National Center for Biomedical Communication, suggested NLM should organize a workshop "to explore what a three-dimensional digital data set of human cadaver anatomy images might be used for [12]". In writing the report of that workshop, Masys gave it the name, "Visible Human Project," which Lindberg then suggested to Ackerman. Lindberg agreed that it should include images from male and female cadavers.

The NLM's request for proposals to carry out the project specified that CT scan, MRI, and anatomical cross-sectional images should be aligned and cross-correlated. The project was awarded to Drs. David Whitlock and Victor Spitzer at the University of Colorado in Denver. The total project cost was $1,400,000, and the resultant image dataset size was 40 gigabytes. Ackerman concludes: "The Visible Human Project was a complete success. It was used in high school, college, and medical school teaching; imaging algorithm development, testing, and comparison; physiological and radiation modeling; art; and digital network testing [12]".

5. Conclusion

This summary has described the wide variety of informatics-related contributions of the NLM and Dr. Donald A. B. Lindberg over a period of 31 years - coalescing to produce an impact on health and biomedicine that continues to this day. The chapter also emphasizes the transformative nature of those contributions, occurring in parallel and synergy with other remarkable changes that include the technological evolution in computers and communication, cultural changes in academic medicine and the health professions, and a new integration of biomedical science with both clinical care and population health.

Much of what observers, clinicians, public health workers, and biomedical scientists now take for granted can be attributed at least in part to the innovations engendered by the NLM, either through its intramural programs or its extramural support of informatics-related research and education.

Furthermore, the contributions have touched the global community by offering free access to the biomedical and clinical literature and to a variety of databases that play key roles in care, prevention, and research. The NLM has supported innovative projects that have created new methodologies that, in turn, have been generalized not only beyond a specific motivating project - but often beyond the biomedical and clinical sciences to other diverse areas of application. Having surveyed the accomplishments and transformative impacts documented herein, the editors and chapter authors stand in awe of what Dr. Lindberg and the NLM accomplished. Humankind owes a great debt to both the man and to the organization he led.

References

[1] L.C. Kingsland 3rd and C.A. Kulikowski, A scientific mind embraces medicine: Donald Lindberg's education and early career. 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), Amsterdam, IOS Press, 2021.

[2] J.H. van Bemmel, M.J. Ball, E.H. Shortliffe and A.B. Donald, Lindberg, pioneer in biomedical and health informatics: his involvement in creating professional organizations. 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), Amsterdam, IOS Press, 2021.

[3] N.M. Lorenzi and W.W. Stead, NLM and the IAIMS Initiative: Cross-institutional academic/advanced systems contributing to the evolution of networked information and resources. 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), Amsterdam, IOS Press, 2021.

[4] R.A. Greenes, V. Florance and R.A. Miller, Don Lindberg's influence on future generations: the U.S. National Library of Medicine's biomedical informatics research training programs. 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), Amsterdam, IOS Press, 2021.

[5] 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), Amsterdam, IOS Press, 2021.

[6] T.T. Kuo and L. Ohno-Machado, NLM's sponsorship of research in biomedical informatics (1985–2016). 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), Amsterdam, IOS Press, 2021.

[7] H. Mo and J.C. Denny, The U.S. national library of medicine's impact on precision and genomic medicine. 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), Amsterdam, IOS Press, 2021.

[8] C.J. McDonald and B.L. Humphreys, The U.S. national library of medicine and standards for electronic health records: One thing led to another. 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), Amsterdam, IOS Press, 2021.

[9] B.L. Humphreys and M.S. Tuttle, Something new and different: The unified medical language system. 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), Amsterdam, IOS Press, 2021.

[10] 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), Amsterdam, IOS Press, 2021.

[11] M.J. Ackerman, S.E. Howe and D.R. Masys, Don Lindberg, high performance computing and communications, and telemedicine. 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), Amsterdam, IOS Press, 2021.

[12] M.J. Ackerman, The visible human project. 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), Amsterdam, IOS Press, 2021.

[13] M.F. Collen, *A History of Medical Informatics in the United States*. American Medical Informatics Association, Indianapolis, 1995.

[14] R.A. Greenes and E.H. Shortliffe, Medical informatics: An emerging academic discipline and institutional priority, *JAMA* **263**(8) (1990), 1114–1120.

[15] National Library of Medicine. Long range plan/report of the Board of Regents. [Internet] Bethesda, Md.: U.S. Dept. of Health and Human Services, National Institutes of Health, January 1987. [cited 2021 Oct 3]. Available from: http://resource.nlm.nih.gov/101646837.