Informatics as Science*

Edward H. Shortliffe
Department of Biomedical Informatics, Columbia University, New York, NY, USA

Summary
The evolution of the informatics field, now with a well-accepted and crucial role in modern biomedicine and health care delivery, is the result of creative research over seven decades. The success is due in part to recognition that, throughout the process, investigators have documented not only what they have done but what they have learned, stimulating and guiding the next generation of projects. Such iterative experimentation, learning, sharing, and progressing is typical of all scientific disciplines. Yet progress depends on identifying key lessons, insights, and methods so that others can use them. This paper addresses the nature of scientific progress in informatics, recognizing that while the field is motivated by applications that can improve biomedicine and health, the scientific underpinnings must be identified and shared with others if the field is to progress optimally.

Keywords
Medical informatics; biomedical informatics; health informatics; scientific research; informatics applications

Yearb Med Inform 2022:303-6
http://dx.doi.org/10.1055/s-0042-1742541

1 Considering the Role of Science in Informatics

It is sometimes possible to identify singular events that mold one's professional career, both practically and philosophically. One such personal event warrants sharing with informatics[1] colleagues since it has influenced me greatly and offers some insights and approaches that may be useful to others.

In 1979, after a decade of medical and informatics training that involved doctoral dissertation research and internal medicine residency, I joined the clinical faculty at Stanford University (Palo Alto, California) as an assistant professor of medicine. Combining both clinical practice and informatics research, I sought to develop an impactful career as an investigator, practitioner, and educator in academic medicine.

Within my first year on the faculty, as I built my research program, I was invited to meet with a senior faculty member in my department. It was an informal collegial meeting – not part of any formal review or advising process. He wanted to give me some advice about how to excel and succeed in the highly rigorous and competitive research environment that characterized Stanford's medical school. He pointed out that, although I had a research program and had already received some grant support, my research was atypical for the medical school and poorly understood by my colleagues. "There is an assumption and bias that biomedical research involves wet-bench labs in pursuit of new basic biological or clinical insights," he pointed out. "If you want to make it at Stanford, and to be promoted to associate professor in time, you will need to convince the senior faculty in the school that what you are doing is science – not just computer programming."

This was sobering advice, particularly because I had assumed that what I was doing *was* science and naively felt that it would be recognized as such by others. It was clear to me that my teachers and colleagues in computer science were not "just" doing computer programming. They were some of the smartest, most innovative, and inspiring investigative minds that I had encountered. Yet the foreign nature of what they did (as viewed from a medical school 40 years ago) required that we in the informatics community make it clear that our work, carefully pursued and presented, constitutes both scientific and applied contributions, just as work in other respected biomedical research fields does.

The advice forced me to contemplate the nature of the science in our field, since the new knowledge that we were creating did not generally involve discovering new biological or clinical phenomena. There is an extensive literature on the nature of science, reflecting a variety of historical and philosophical perspectives. Analyses tend to emphasize discovery, innovation, reproducibility, testability, and empiricism [1, 2]. While motivated by applied goals (as the results of other medical research ultimately are), the new knowledge offered by informatics researchers primarily focuses on methodological insights and innovations,

* This paper is adapted from a presentation given by the author at the International Medical Informatics Association's François Grémy Award of Excellence ceremony during Medinfo2021, October 2, 2021 (held virtually).

[1] This paper uses the term informatics as the generic name for our discipline, without an adjective. The field has various names, depending on local or national customs, so the term informatics here is intended to encompass biomedical informatics, medical informatics, health informatics, and similar naming conventions.

coupled with discovery of explanations for observed phenomena that has been called a "local science of design" [3].

With the development of our Stanford graduate training program in biomedical informatics, which we founded in 1982, we realized that we had to train our students to be informatics scientists. Students were understandably inspired by a desire to address an applied problem that they had identified in the real world of biomedicine, using computational or related methods, but they had to be taught to recognize that the end-product application was generally not a scientific contribution until it had been analyzed and shared with others. We required students to ask, "How does my work contribute to the field of informatics?" rather than simply, "Does the contribution successfully meet its applied goals?" A contribution to the field offers methods or insights that generalize so that others can use them in their own work. Our students must recognize the cyclical nature of science, with each effort ideally contributing to, and elevating, the subsequent research by ourselves or others who work in the same area.

2 An Approach to Pursuit of Informatics Scientific Goals

The evolution of my thoughts regarding a methodological and scientific approach to informatics research has resulted in the creation of a generic multi-step process that captures what I try to convey to students and research colleagues (see Box 1). It is intended to offer a structure that individuals can consider and adapt as is suitable for their own research and development work in our field. There is an initial series of steps that precedes the actual hands-on research effort ("Before"). Then work on the project itself has several steps that need to be undertaken ("During"). What follows in the third stage is the analysis and sharing of results – a crucial part of the scientific endeavor ("After").

Note that the ordering of individual steps is intended to be flexible and may need to be adjusted depending on the specifics of the individual project (e.g., goals, size, and where it lies on the basic-to-applied spectrum). In addition, it may be appropriate to publish interim results (and the associated lessons or methods) while the overall project continues toward completion. Even the most applied projects, which at first may appear to be using standard methods and

Box 1 Scientific Steps in The Evolution of an Informatics Research and Development Effort (see text)

Before: (defining and designing the research effort)	**Identify**	Ask: "What is the problem to be addressed?" "Has it been addressed before?" "Successfully?" Clearly state the answers to these questions. Seek consensus regarding the research idea with other interested parties (e.g., colleagues, advisors, potential collaborators).
	Partner	Assess the expertise of the existing team. Define the kinds of collaborators who will be required, especially those from the application domain. Recruit multidisciplinary team members as appropriate. Be prepared to seek others as plans evolve and missing expertise becomes clearer.
	Analyze	Dig deeper into the problem. Read pertinent background papers or books. Consider methodological options and, if appropriate, theoretical rationale for your approach. Aim to become an expert on the topic to be addressed and the approaches that could be selected and applied.
	Motivate	Identify and engage collaborators. Aim to excite them with the problem selected and the proposed approaches to be considered. Seek to develop a team that is eager to be involved and positive about the potential value of the end results.
	Create	Develop a plan to proceed with the work. Attract the necessary required resources (fiscal, data, technology, etc.), recognizing that some may not become available until preliminary work has been undertaken. Identify what is unique about the approach envisioned.
During: (executing the plan and assessing results)	**Innovate**	Seek to identify at the outset how the work could add to informatics knowledge, recognizing the cyclical nature of science that requires experimentation, feedback, rethinking, and trying again. Anticipate that such insights will evolve as the work proceeds.
	Implement	If system building, begin the construction and testing. Acquire pertinent data and use them to guide what you do as the work evolves. Undertake analytical experiments as required (formative evaluations). Be prepared to encounter problems that require redirections. Engage stakeholders as appropriate.
	Assess	Design and carry out summative evaluations of the work once the team believes it is ready. Such experiments are more formal than those formative efforts during the evolution of the implementation. Assure rigor so that result is meaningful and convincing to others.
After: (reflecting on results and communicating)	**Generalize**	While assessing results, ask how they will be useful to others. Try to identify other problems or domains that are amenable to being approached using the same methods that have been developed. Your scientific contribution will follow from the generalizable lessons.
	Critique	Assess weaknesses or limitations. Consider how they should be addressed in future work by yourself or others. Ask whether the assessment should simply be redesigned or if the problem lies with the larger system or approach. Recall that negative results offer insights that can usefully be analyzed and shared.
	Share	Present work in papers and presentations, emphasizing the generalizable results. If a new method was developed, convey its range of applicability. Refer informatics science to suitable informatics journals or meetings. Send papers on application results to journals in that domain.
	Inspire	Give well-prepared talks that excite your audience. Recognize that even negative results can be useful to others when effectively analyzed and shared. The lessons of your work can inform the subsequent work of others. In fact, that is your primary goal!

approaches, often will result in generalizable insights or lessons that are worthy of sharing with others[2]. The science of our field demands that we identify and share not only what we have done but also what we have learned. Papers, books, dissertations, talks, and other types of presentations will advance the field if those lessons are convincingly identified, supported by data, and effectively communicated.

3 Implementing and Sharing the Perspective

The philosophy and guidance outlined here have guided my own work as well as my education of others. I first wrote and spoke about the topic at a meeting organized by an IMIA working group in Chamonix, France in 1983 [4]. Subsequently, our students in the Stanford informatics training program were taught to identify the generalizable scientific lessons in their own work and to write or give talks about them accordingly.

Then, in 2000, I oversaw the rebirth of the former *Computers and Biomedical Research* (Academic Press) when I became editor of the *Journal of Biomedical Informatics*. We designed the journal to focus on papers that stressed the scientific results of informatics research, briefly describing our editorial goals as follows in the inaugural issue:

> *The Journal of Biomedical Informatics (JBI) is intended to complement rather than to compete with the other major journals in medical informatics. In particular, we wish to emphasize papers that elucidate methodologies that generalize across biomedical domains and that help to form the scientific basis for the field. Papers will tend to be concerned with information technology rather than medical devices, and on underlying methods rather than system descriptions or summative evaluations* [5].

Now published by Elsevier (which acquired Academic Press shortly after the journal was introduced under its new name), JBI has become known as "the premier methodology journal in the field" as it continues to emphasize the publication of papers that advance the underlying science of the informatics discipline [6].

More recently, I began to consider how to encourage students from other informatics training programs to assess and convey effectively the scientific content of their work. I was concerned that many doctoral dissertations in our field were more focused on a specific application and its description than on the underlying novelty and scientific contributions of their work. One way to increase awareness of my concern was to publish a paper about the philosophy, goals, and content of informatics PhD dissertations [7]. Thereafter I joined with others to propose and fund an annual doctoral dissertation award to be bestowed by the American Medical Informatics Association (AMIA). As is noted on their web site, "The AMIA Doctoral Dissertation Award offers high-value and prestigious recognition for the top doctoral dissertation each year that contributes to the science of informatics in any biomedical application domain or domains" [8]. The award has been given annually since 2017 and the finalists' dissertations are available on the AMIA website for review by prospective nominees, thereby inspiring their own doctoral work and the way that they may choose to write about it.

The recommendations discussed in this article are in no way meant to decrease the importance or impact of the impressive and novel informatics applications that advance biomedicine, clinical care, and population health. Everyone in the field is motivated by a desire to have a positive impact to address the scourge of disease or to promote public health. Rather the goal here is to recognize that almost every project discovers new truths, new methods, or new ways of thinking about problems. It is the responsibility of those who do the work and ultimately share it with others to identify the useful innovations and lessons, emphasizing their range of applicability plus their strengths and limitations, so that the scientific base of informatics is advanced.

Acknowledgments
I am grateful for the passion, talents, accomplishments, and friendship of many informatics students and colleagues over the years. In addition, much of the philosophy and research approach summarized in this article was influenced by the work and advice of two of my early mentors – both eminent scientists known largely for work beyond the field of informatics. Joshua Lederberg, a Nobel Laureate and geneticist [9], served as a superb role model for me as he conceived and directed, beginning in 1973, the NIH-funded *Stanford University Medical Experimental Computer for Artificial Intelligence in Medicine* (SUMEX-AIM) [10]. Stanley N. Cohen, a physician and geneticist, pursued early work on computer-based drug-interaction warnings at Stanford Medical School [11], a project that involved me during my first year in medical school. Subsequently Cohen went on to serve as my PhD advisor. His most influential work is largely in genetics [12], but he also was an exemplar of the rigorous approach to scientific research, adapting what he knew to informatics problems as he guided the author's dissertation work.

References

1. Heilbron JL, editor. The Oxford Companion to the History of Modern Science. 1st edition. Oxford; New York: Oxford University Press; 2003.
2. Lindberg DC. The Beginnings of Western Science: The European Scientific Tradition in Philosophical, Religious, and Institutional Context, Prehistory to A.D. 1450. 2nd edition. Chicago: University of Chicago Press; 2008.
3. Patel VL, Kaufman DR. Science and practice: a case for medical informatics as a local science of design. J Am Med Inform Assoc 1998 Nov-Dec;5(6):489-92.
4. Shortliffe EH. The science of biomedical computing (keynote address). In: Pages J, Levy A, Gremy, F, Anderson J, editors. Meeting the Challenge: Informatics and Medical Education [Internet]. Amsterdam: North Holland; 1983 [cited 2022 May 11]. p. 1–10. Available from: https://doi.org/10.3109/14639238409015189
5. Shortliffe EH. Editorial: Inaugural Issue of JBI. J Biomed Inform 2001 Feb 1;34(1):2–3.
6. Journal of Biomedical Informatics (Elsevier) [Internet]. Premier methodology journal in the field. https://www.sciencedirect.com/journal/journal-of-biomedical-informatics
7. Shortliffe EH. The organization and content of informatics doctoral dissertations. J Am Med

---

[2] Note that this is true in industrial settings as well as in academia. Commercial efforts often have much to offer to the underlying science, but the developers will overlook this important step if they focus solely on the product and its commercial viability.

Inform Assoc 2016 Jul;23(4):840-3.
8. AMIA Doctoral Dissertation Award [Internet]. AMIA - American Medical Informatics Association. [cited 2022 May 11]. Available from: https://amia.org/about-amia/amia-awards/research-awards/amia-doctoral-dissertation-award
9. National Library of Medicine. Joshua Lederberg: Profiles in Science [Internet]. [cited 2022 May 10]. Available from: https://profiles.nlm.nih.gov/spotlight/bb
10. Shortliffe EH, Rindfleisch TC. Presentation of the Morris F. Collen Award to Joshua Lederberg, PhD. J Am Med Inform Assoc 2000 May-Jun;7(3):326-32.
11. Tatro DS, Briggs RL, Chavez-Pardo R, Feinberg LS, Hannigan JF, Moore TN, et al. Detection and prevention of drug interactions utilizing an on-line computer system. Drug Inf J 1975 Jan-Apr;9(1):10-7.
12. National Library of Medicine. Stanley N. Cohen Papers: Open for Research [Internet]. Circulating Now from NLM. 2019 [cited 2022 May 10]. Available from: https://circulatingnow.nlm.nih.gov/2019/01/24/stanley-n-cohen-papers-open-for-research/

Correspondence to:
Edward H. Shortliffe, MD, PhD
272 W 107th St #5B
New York, NY 10025, USA
Tel: +1 212 666 8440
E-mail: ted@shortliffe.net
https://www.shortliffe.net