The Labor Market for Individuals with Advanced Degrees in Biomedical Informatics

Sheila Dugan
Jason Saunders

The Thomas Jefferson Program in Public Policy

The College of William & Mary

December 15, 2011

Table of Contents

Tables and Charts

ACKNOWLEDGEMENTS

The team would like to take this opportunity to thank all those who helped contribute to this report. Thank you to Dr. Valerie Florance at the National Library of Medicine for her input and support, as well as Professors Robert Archibald and Lou Rossiter for their guidance and suggestions. Thank you also to Ankita Aras for her assistance in the research and data analysis.

A special thanks to those who were able to provide us with key information and background in the biomedical informatics field including Dr. Don Detmer of the University of Virginia and Dr. Lawrence Hunter of the University of Colorado – Denver.

The opinions represented here are those of the authors and do not necessarily represent the views of the National Library of Medicine or National Institutes of Health.

Executive Summary

Our research found that there was sufficient demand for individuals with advanced degrees in biomedical informatics. A review of the job advertisements in *Science* magazine revealed a decline in the number of relevant advertisements printed over an eight year period. On the other hand, the online source, Bioinformatics.org, saw an increase in the number of job postings for individuals with degrees in this field over the course of four years. Even with the observed decrease in the print source, we estimate that the number of jobs each year exceeded the number of graduates produced by biomedical informatics programs.

The survey distributed to degree recipients in this field revealed graduates had no trouble finding jobs. Most of our survey participants indicated they had a job upon graduation and the longest job search lasted seven to twelve months. The most common method of finding employment was through networking. Most graduates entered the academic sector upon graduation, with the heaviest concentration found in academic postdoctoral positions. A majority of survey respondents indicated that research was one of their primary responsibilities at their initial job. The individuals with the lowest reported starting salaries were postdoctoral research associates or assistants. Only five individuals said they experienced a long period of unemployment.

Given the analysis of job postings and trainees, as well as the labor market experiences reported by graduates, there does not seem to be a need to cap the number of individuals seeking advanced degrees in biomedical informatics. Demand for these trainees appears to be strong and there is no indication of a surplus of trainees.

Introduction

A current major concern of the National Institutes of Health (NIH) is the future of the biomedical research workforce. NIH and other U.S. government agencies such as the National Academy of Sciences and the National Research Council have all established committees to examine the needs and challenges facing biomedical researchers. The National Library of Medicine (NLM) is interested in the research labor market as well, due to their almost 40 years long investment in training programs for biomedical informatics. This study presents research about the labor market for individuals with advanced degrees in the field, including information about the supply of trainees and the how well they fare once they enter the market, which will aid NLM in future decisions about funding trainees.

The report will first discuss the background of NLM's support of the field. Next, similar past studies will be analyzed. The following sections contain quantitative analyses on the number of trainees and the demand for trainees, as well as the results of a survey of past trainees. Finally, the major implications of the report will be explained along with recommendations for next steps and future studies.

Background

The National Library of Medicine (NLM) currently supports pre- and post- doctoral biomedical informatics training programs at eighteen colleges and universities. NLM funds biomedical informatics training programs at Columbia University, Harvard University, Indiana University, Johns Hopkins University, Oregon Health and Science University, Rice University, Stanford University, University of California – Irvine, University of California – Los Angeles, University of Colorado, University of Missouri – Columbia, University of Pittsburgh, University of Utah, University of Virginia, University of Washington, University of Wisconsin – Madison, Vanderbilt University, and Yale University (The National Library of Medicine 2011).

Public and private institutions of higher education are encouraged to apply for the training grants. The purpose of the grants is to support students while they undergo their training. Individuals without a doctoral degree must leave the program with a PhD in biomedical informatics or a related field. If one enters the training programs after attending medical school or receiving a doctorate degree in another field, they must receive a master's degree that allows them to conduct research or another PhD. The goal of the five-year grants is to train a generation of researchers to contribute knowledge to the field and develop new informatics tools and technologies. Graduates of the program are not expected to pursue administrative roles in hospitals and other healthcare institutions.

NLM identifies four major informatics training areas in its grant application instructions. Healthcare/clinical informatics use informatics tools to care for patients. Translational bioinformatics promotes transforming research into real-world applications. The clinical research track applies these tools and principles to support clinical trials and research. Public health informatics works towards the development of tools for health services research, decision support in public health agencies and governmental organizations, and to monitor disease outbreaks (The National Library of Medicine 2011).

Between Fiscal Year (FY) 2001 and 2011, NLM spent $166.22 million on graduate training programs in biomedical informatics. Each school received an average of $923, 467 each year. NLM awarded Harvard University the most money, dispersing over $20 million to the medical school in the past ten years (The National Library of Medicine 2011). The table below offers details on the total funding dispersed to each school currently funded since fiscal year 2001:

Table 1: Institutions Receiving National Library of Medicine Funding

Institution	Total Funding (FY 2001-FY 2011)
Columbia University	$17.44 million
Harvard University	$20.48 million
Indiana University/Regenstrief Institute	$4.51 million
Johns Hopkins University	$5.27 million
Oregon Health and Science University	$7.28 million
Rice University	$8.32 million
Stanford University	$11.07 million
University of California Irvine	$9.15 million
University of California Los Angeles	$5.82 million
University of Colorado Denver	$3.07 million
University of Missouri Columbia	$5.85 million
University of Pittsburgh	$13.49 million
University of Utah	$10.07 million
University of Virginia	$3.92 million
University of Washington	$9.46 million
University of Wisconsin	$11.18 million
Vanderbilt University	$8.94 million
Yale University	$10.90 million

Many schools offer advanced degrees in biomedical informatics that do not receive funding from NLM. American Medical Informatics Association (AMIA)'s Academic Forum is a consortium of schools offering training in biomedical informatics or related disciplines. Programs not funded by NLM grants, but are members of the Academic Forum include: Arizona State University, University of Miami, and other schools across the United States. Most of the schools offer master's degrees in the field. Some of the schools recently established their informatics training programs. For example, Emory University established its Center for Comprehensive Informatics in September 2008 and launched its Department of Biomedical Informatics in 2011 (Center for Comprehensive Informatics, Emory University 2011).

LITERATURE REVIEW

In an article published in the International Medical Informatics Association (IMIA)'s 2008 Yearbook, Dr. William Hersh at Oregon Health and Science University admitted very little was known about informaticians in the medical field. Hersh wrote, "...both healthcare leaders and informatics leaders need more information upon which to base implementation of systems, optimal deployment of the workforce, and the best educational options for the workforce" (Hersh, William 2008, p. 162). To properly plan for the expected growth in this field, which may include the need for more individuals to be trained, more information must be known about the composition of labor force, the rate of growth, and the types of companies, organizations and institutions who demand the services of informaticians.

Five major studies were identified that focused on the labor market for individuals with advanced degrees in biomedical informatics. These studies failed to fully answer the research questions proposed in this report. The number of years they studied the labor market was too short, or the researchers restricted their analysis to a few schools. A study of two German universities is difficult to apply to the American market because there is no indication that these two markets are similar. Despite these shortcomings, a review of previous studies can be used as a source for information on the labor market, as well as provide a framework for analysis on this topic.

Medical Informatics Education published a report on the graduates of the University of Utah's medical informatics program in 1999. (The University of Utah is one of the programs currently receiving an NLM research training grant.) Gregory A. Patton and Reed M. Gardner, the study's authors, used university records in an attempt to identify every individual who graduated from the program since 1964. They excluded from the study non-degree seeking individuals who were at the university to conduct research. Patton and Gardner then located and contacted the Utah alumni via telephone, e-mail, or in-person.

Patton and Gardner convinced 272 current and former graduate students to participate in their study. Among the students in the study, 205 came from the United States, while 67 students were from other countries. Ninety-six students entered the program after receiving one to three prior graduate degrees. Fifty-seven individuals had already received a medical degree before beginning the program. The average length of time it took to complete a master's degree was 3.1 years. It took an average of 4.5 years to complete a doctoral degree.

The University of Utah study found that nine students entered medical school after the program. Industry was the most popular type of organization that employed former students, accounting for the jobs of 37 percent of those surveyed. Integrated delivery system organizations employed 27 percent of the graduates. Twenty-three percent were employed at an educational institution and 6 percent worked in a medical practice. Only 18 alumni took a job outside of the country, with most of the alumni living and working in the United States. Twenty-six students were without jobs at the time of the study, with 24 of those students being current students. Patton and Gardner's research revealed the graduates of Utah's informatics programs took jobs in private industry and provided no evidence that graduates of the program had trouble finding jobs (Patton, Gregory A. 1999).

P. Knaup, W. Frey and R. Haux conducted a similar study for individuals earning the equivalent of a master's degree from the Universities of Heidelberg and Heilbronn's medical informatics program in Germany. The purpose of the study was to learn more about the employment prospects and job histories of individuals who graduated. The researchers distributed a survey to the first 1,024 graduates of these two programs. They received responses from 446 individuals.

Knaup et al.'s study revealed 43 percent of graduates worked in the medical informatics field, while 51.4 percent work in an informatics field outside of medicine. Most individuals not working in the medical informatics field blamed the “labor market” and “personal reasons.” Graduates who worked for hardware/software companies made up 32.9 percent of the sample and 19.4 percent worked for other types of private companies, making the private sector the primary employer of medical informatics graduates from these two schools.

The researchers asked the survey participants to select which area they are working in or the type of tasks that take up at least 10 percent of their working day. The majority of the participants worked on information systems in healthcare and medical documentation. The gross annual income for 37.3 percent percentage of the graduates was between 50,500 to 75,000 Euros and 34.4 percent of graduates reported they made between 25, 500 to 50,000 Euros each year. Only one individual made less than 5,000 Euros.

When asked about the first three job changes, the most popular answer (38.4 percent) reported that their job changes were within the medical informatics field. Less than one percent of the graduates reported moving out of the medical informatics or general informatics field. A little over 75 percent of the survey participants described their job situation as “very satisfactory” or “satisfactory (Knaup P 2003).”

Grant C. Black and Paula E. Stephan have studied the bioinformatics field extensively. In January 1998, Black and Stephan presented a paper at a workshop sponsored by the National Research Council's Committee on Science, Engineering, and Public Policy on the topic. The paper would later be published in *Science and Public Policy*. The authors of the study examined the number of new job postings in *Science* magazine published in 1996 and 1997 to determine how many jobs were available for individuals with training in biomedical informatics. These are considered lower-bound estimates. Positions were only counted if they mentioned computational biology or bioinformatics in the advertisements. Some advertisements did not contain details on the number of positions offered by the business or organization, making it difficult to discern exactly how many positions were being advertised each year (Black, Grant C. 1999, p. 1).

Black and Stephan observed 209 positions advertised in 1996 and 354 positions advertised in 1997 (Black, Grant C. 1999, p. 4). The researchers found that the majority of advertisements were placed by firms, with only 36 universities placing advertisements in the two years. Entry-level positions, as well as senior-level positions for scientists, were advertised in the magazine. Most positions, however, required that individuals have a doctorate degree (Black, Grant C. 1999, p. 4).

In May 1999, Black and Stephan examined hiring patterns for individuals with degrees in bioinformatics and computational biology for the Alfred P. Sloan Foundation. The authors surveyed formal training programs. The survey incorporated questions on the training program, including the degrees awarded, the number of students in the program, and what financial resources they received to run the program.

The primary purpose of the survey was to discover what type of institutions were hiring graduates with undergraduate, masters, and PhD degrees in bioinformatics and computational biology. The survey asked participants to name hiring institutions and the number of students placed at that particular organization. Black and Stephan collected a range for the starting salary, as well as information on how students discovered job opportunities (Black, Grant C. 1999).

The results of the survey repudiated the idea that the services of the informatics students was in so much demand that individuals left the training programs before they received their degrees. Only seven of the fifty-three jobs were at academic institutions, and only one job was found at a governmental organization. Most students took jobs in the private sector at biotechnology or pharmaceutical firms (Black, Grant C. 1999, p. 3). Stephan and Black observed that salaries increased with more training. Individuals with a PhD in the field reported making over $70,000 in their first job in the field (Black, Grant C. 1999, p. 4). The most common method used to find out about job opportunities were print advertisements, validating Stephan and Black's earlier work that estimated demand by counting the number of advertisements in *Science* magazine.

Stephan and Black published a more detailed study of the field in *Biotechnology Education* in 2005. According to their research the field changed dramatically since the 1990s (Black, Grant C. 2005). The number of programs offering training in bioinformatics and computational biology and the number of graduates of these programs increased. Stephan and Black utilized the same method to estimate demand: counting the number of positions posted in *Science* magazine in 2000, 2001, and 2002. This time however they discovered the number of advertisements decreased each year. There was also a shift in the types of organizations that posted advertisements in the magazine. Between 1996 and 1997, most of the job advertisements were posted by businesses. Academic institutions posted most of the advertisements in the new years Black and Stephan studied (Black, Grant C. 2005, p. 61). The authors found these two trends to be troubling, and predicted that graduates would have trouble finding jobs if these trends continued.

The few labor market studies conducted about individuals with advanced degrees in biomedical informatics focused on where the graduates were placed after their training programs. Patton and Gardner's study only examined the students that attended the program at the University of Utah, while Knaup et al. studied two universities in Germany. These two studies did not encompass a diverse enough group of biomedical informatics graduates to draw conclusions on the state of the labor market. Stephan and Black's reach was more extensive. In their early research in the late nineties, Stephan and Black observed growth in the field, but were concerned about the decreasing number of jobs advertised in *Science* magazine by 2002.

They were also concerned about the shift in the types of institutions seeking employers. In the 1990s more private companies posted job advertisements. By 2001, academic institutions posted more advertisements in *Science* magazine, indicating a major change took place in this market.

Table 2: Summary of Literature Review

Study	Years Examined (Job Market)	Sample size (Number of universities)	Findings
Stephan and Black; "Bioinformatics: Does the US System Lead to Missed Opportunities in Emerging Fields?" (1999)	Two years (1996 and 1997) of job postings in Science magazine were examined.	N/A	The educational system is not responding to the increased demand for bioinformatics graduates.
Stephan and Black; "Hiring Patterns Experienced by Students Enrolled in Bioinformatics/Computational Biology Programs" (1999)	N/A	Sixteen academic institutions offering undergraduate, graduate or postdoctoral training in bioinformatics participated in the survey.	The number of individuals graduating from formal bioinformatics training programs are not enough to fill the available jobs.
Stephan and Black, "Bioinformatics Training Programs are Hot but the Labor Market Is Not" (2005)	Five years (1996-1997; 2000-2002) of job postings in Science magazine were examined.	Forty-four academic institutions offering undergraduate, graduate or postdoctoral training in bioinformatics responded to part or all of the survey distributed.	Demand in the field has declined, while the number of bioinformatics programs increased. The major employers are now academic institutions as opposed to private sector companies and organizations.
Patton and Gardner; "Medical Informatics Education: The University of Utah Experience" (1999)	N/A	One academic institution, the University of Utah's graduate program in medical informatics, was studied.	Most graduates work in health care delivery organizations. 91 percent of alumni were employed in the United States.
Knaup, et al.; "Medical Informatics Specialists: What Are their Job Profiles?" (2003)	N/A	One joint medical informatics graduate program being run by the University of Heidelberg and the University of Applied Sciences Heilbronn.	One-third of graduates surveyed are working in software/hardware companies. The variety of jobs and profiles demonstrate programs like the Universities of Heidelberg and Heilbronn

The method employed by Stephan and Black to approximate demand is not unique. Michael D. Basil and Debra Z. Basil studied the market for marketing professors at business schools. Basil and Basil used two methods to determine demand for marketing professors. They first used online archives to examine every job posted between 1995 and 2003. Each position originating in the United States was then labeled based on whether it was a permanent, tenure track or temporary position, rank, and the specialty. The researchers also measured demand by looking at the number of employers registering for the Academic Placement Service at the Southern Marketing Association. The increase in the number of employers looking for individuals to hire means that the services for marketing professors are highly demanded (Basil, Michael D. 2006).

Marybeth F. Grimes and Paul W. Grimes examined the labor market for individuals with master's degrees in library science to determine whether there was a demand for their services. Grimes and Grimes collected data on job advertisements published in *College and Research Library News*. Only seven years were included in the sample: 1975, 1980, 1985, 1990, 1995, 2000, and 2005. The information was used to create a database of over 4,000 records (Grimes, Marybeth F. 2008, p. 333).

Overall, researchers used two methods to examine the labor market. They distributed a survey to key institutions or the populations being studied. In addition to developing and disseminating a survey, they reviewed job advertisements in the field to see the number of available positions. These methods were utilized in the labor market study for individuals with advanced degrees in biomedical informatics.

Primary Research Question And Methodology

The major area of interest to the National Library of Medicine is the labor market for these informatics-trained researchers. Given the potential job openings, what is the "right" number of trainees? Should there be an effort to cap the number of trainees? It is beyond the scope of this study to calculate a specific number of individuals given the convenience sample of individuals included here, but by examining trends over time during the past 10 years in job openings and trainees, approximations for portions of demand and supply can each be derived. The analysis of these estimates can provide some insight into whether there is a surplus of qualified trainees, in which case there should be a cap on new trainees, or whether there is a shortage of trainees, which would not require a cap.

In addition to the general labor market, this report will provide information about two areas specific to NLM's support of the field. First, we will use data on trainees who received support from the training grants over the past 10 years and compare these aggregate counts to the estimates for total number of graduates. This will allow NLM to see how much impact its grants have on the number of trainees. Second, using results from a survey, we will be able to show how many past trainees found jobs involving research, one of the key objectives of the biomedical informatics training grant program.

Assessing Supply

The supply side in a labor market study is comprised of the workers participating in the field. We used two methods to gain a better understanding of the supply side of the market. First, we established a list of identified programs. The list included the 18 NLM-funded programs, but we recognized these universities did not represent the total market for new PhD trainees. Since several of the NLM-funded programs belong to AMIA's Academic Forum, we used the list of Academic Forum members as a starting point for identifying other PhD programs. By visiting the websites of each non-NLM-funded program, we identified 9 other PhD granting institutions. It should be noted, however, that there was less uniformity in the degrees offered by the non-funded programs. Some offered doctoral degrees in nursing with a specialty in informatics, while others offered biomedical informatics as a specialty of either a general informatics degree or a degree such as computer science.

Once we established the list of university programs, we revisited each program's website to search for lists of alumni and current students. Nearly all of the NLM-funded programs published a list of either current students, alumni, or both while only 3 of the 9 non-NLM programs had placed such lists on their websites. There were some challenges in interpreting these lists, especially in regards to alumni. Some programs provided a clearly delineated list of alumni organized by pre- and postdoctoral participants, the year the degree was received, and whether that degree was a doctoral or master's. Other programs offered more simplified lists that did not include one or more of these organizing characteristics. The crucial element for this project was whether the program identified degrees by the year awarded. In order to keep the methodology consistent and not inflate supply numbers, only those programs which had alumni that graduated between 2001 and 2011 easily identifiable were included in the analysis. Efforts were also made to only include those students who were seeking doctoral degrees or who were postdoctoral students. Similar efforts were made when examining the lists of current students. After our restrictions were put in place, our list of programs only included NLM-funded programs[1].

There a few limitations on the supply side analysis. First, since not every program publishes a list of alumni or current students, or if they do the list did not meet the required

[1] The programs that met our criteria are listed in Table 3.

criteria, any estimates of workers in the field represent lower bound estimates. Next, some individuals may request that their names not be published on the programs website for privacy concerns, which will also push down estimates. Related to this, some programs have given funding to individuals who either were not degree seeking or who left the program completely. These individuals will most likely not be listed on the program's website. Another potential concern is using Academic Forum membership as a starting point for identifying other training programs. Several, but not all, of the NLM-funded programs are Academic Forum members. Other non-NLM-funded programs may also not be members, which would lower trainee estimates. Finally, given the interdisciplinary nature of the field, there may be instances where doctoral graduates of related programs, such as computer science, receive some type of medical or clinical training that would give them the skill set necessary to enter the biomedical informatics labor market. These individuals would be seeking the same jobs and positions, but are not readily identifiable.

Assessing Demand

Building upon the work of Stephan and Black, a three-pronged approach was taken to approximate demand in the field. First, because of the many benefits explained by Stephan and Black, we collected and analyzed job postings in *Science* magazine. We included only the years 2001 through 2008 due to time limitations. This represents a widening of years of data, however, as past researchers have only done 2 to 3 years of analysis at one time. A broad set of search terms were used when looking through the job postings, including but not limited to: biomedical informatics, bioinformatics, computational biology, biostatistics, or job descriptions that indicated some intersection of the biomedical/life sciences and computer science. We avoided certain terms, such as bioengineering or biochemistry. These fields may be tangentially related to biomedical informatics, but often require a different set of skills or educational background. The research group then narrowed down the collected postings for uniformity. Efforts were made to eliminate duplicate job postings in the same calendar year.

The second approach was to use an online job posting website as a data source. The major challenge of this approach is that many online job sites do not archive postings for longer than 6 months to a year, which limits the ability to do a trend analysis. We utilized the website Bioinformatics.org because it worked around this limitation and allowed searches going back 4 years. It is also helpful in that it is targeted towards individuals in the field of bioinformatics. Similar restrictions and efforts were undertaken with the online postings as were used with the print source.

Estimates may differ from past research for two reasons. First, we attempted to narrow down job postings that fit the search terms, but which also had a biomedical application. This was done to exclude job postings which were posted by employers such as the USDA which sought individuals with a background in bioinformatics for work on the study of crops. Second, when a job posting indicated a definite number of positions, such as "3 Assistant Professor Positions," the job posting was counted 3 times. If the posting simply indicated "multiple postings," it was only counted once.

This portion of the demand side analysis contains several limitations. It relies upon job postings being placed in one of two places: either the magazine *Science* or the website Bioinformatics.org. In reality, many jobs may not be posted in either location. Many jobs may rely on internal hiring or be passed along by word of mouth, which are not accounted for in this method. In counting the advertisements, human error may play a factor due to the large number of job postings each week in *Science*. One final concern is that these two sources may not have sufficient coverage for jobs in the clinical, medical, or public health informatics fields. Other sources, such as AMIA's online career center, may present job advertisements that are more relevant to individuals in those areas. These reasons make the demand estimates lower bound estimates.

Survey

As a final method of examining demand, we created and released a survey to graduates of the NLM-funded programs. The results of the survey provided a firsthand look into the labor market for biomedical informatics trainees. We structured the questions around the experience the trainee had had following graduation, such as how long their job search took, what sector was their initial job in, and how long they stayed at their initial job. The survey asked respondents to provide their name at the beginning, allowing the research group to identify past NLM-funded trainees and examine whether they had different experiences in the labor market than their peers. The survey also asked respondents to identify their educational path (pre-doctoral candidate seeking a PhD, postdoctoral candidate seeking a masters, etc.) in order to allow for potential comparisons between PhD and master's recipients or pre- and postdoctoral trainees.

On October 7, 2011, the NLM program director for extramural programs sent an e-mail of introduction to the program directors of the training programs funded by NLM. The e-mail contained information about the survey and encouraged program directors to participate. Twelve schools agreed to participate in the study. They included Columbia University, Harvard University, Indiana University, Johns Hopkins University, Rice University, Stanford University, University of California – Irvine, University of Colorado, University of Pittsburgh, University of Virginia, University of Washington, and Yale University. In addition, the research team sought the contact information for individuals who attended NLM-funded universities which decided not to participate. The team emailed these individuals directly with a letter of introduction and an invitation to participate in the survey. The research group distributed the survey on October 26, 2011. The survey closed on November 14, 2011.

After the survey closed, we excluded surveys which were not completed. A survey was deemed as not having been completed if the survey software indicated that the respondent had not gone through every question and finished. The only question that respondents were required to answer was the question asking about their educational path. Respondents who only answered that question were also excluded for the purpose of the analysis. After these restrictions were put in place, the team had 116 valid survey responses, 106 of which were from degree-seeking students.

The major limitation of the survey is that it only included NLM-funded programs. An effort was made to include non-NLM-funded members of AMIA's Academic Forum, however this was not possible due to time constraints. The responses are also heavily biased towards NLM-funded trainees with nearly 60 percent of degree-seeking students being identified as individuals who received funding. This figure could actually be higher, as some respondents did not provide a name or received their degree prior to 2001. Access was only given to individual funding level data going back to 2001, which prevented identification of NLM-funded trainees prior to that year. This will be discussed further in the survey results section. Finally, there are the concerns associated with any survey. Respondents choose whether or not to answer the survey. Some who have had less than ideal labor market outcomes may choose to not answer and thus the survey results may paint an overly positive picture.

Some universities had larger amounts of participation than others: There were 28 respondents who identified Stanford University as where they completed their training, 19 who attended Columbia University, 15 who entered Harvard University and/or MIT, 9 from Yale University, 7 from the University of Washington and from Vanderbilt University, and fewer than 5 responses from the following institutions: Indiana University, John Hopkins University, Rice University, University of California-Irvine, University of California-Los Angeles, University of Pittsburgh, University of Utah, and University of Wisconsin. The high number of responses from a small number of universities has the potential to bias the results towards those universities, but when examined individually, responses from a single university do not differ significantly from the sample at large. It should be pointed out that some of these programs are larger than others and thus we would expect them to have higher rates of participation. Furthermore, this list shows that we received a good sampling of the various types of programs that NLM funds and the survey results are not biased towards one of the four particular training areas.

A challenge we faced was calculating an appropriate response rate for the survey. Based upon our investigation into the number of past trainees, described below, the team estimates that approximately 578 pre-doctoral candidates graduated with PhDs from the 18 NLM-funded universities over the past 10 years, and 351 postdoctoral candidates have graduated with advanced degrees over the same time frame, for a total of 929 individuals. We derived these numbers by doubling the numbers of pre-doctoral trainees found at 9 universities and tripling the number of postdoctoral trainees at 6 universities. These numbers are most likely over-estimates, as the method assumes that universities have graduated similar numbers of trainees over 10 years. Several of the NLM-funded programs did not exist 10 years ago, which would lower these estimates. Even with this difficulty, this is the best option available to the team without a list of past trainees from each university. Using these estimates, the survey had an 11.4 percent response rate.

Assessing Supply: The Number of Trainees

Using the lists of current students and alumni available, a better understanding was gained of the number of individuals entering the labor force. In analyzing trainees, we made an effort to keep pre- and postdoctoral students separate. Data was also provided to us by the National Library of Medicine which indicated NLM-funded trainees by the program they attended and the year they earned their degree. This information will be useful in determining the degree of support that NLM provides to the universities. It should be noted that NLM training support can only go to permanent residents of the United States. We were unable to determine from the lists we gathered which students were permanent residents and which were international students. This fact could push down any estimates for the number of students that receive NLM funding.

Pre-doctoral Candidates

Nine programs provided lists of alumni and 11 programs provided lists of current students who met the research group's methodological specifications. From the 9 alumni lists, we estimated a total of 289 graduates over the past 10 years. National Library of Medicine training grants supported 174 trainees at these 9 programs, meaning that in the aggregate, NLM funding supported over 60 percent of these trainees. Among the universities that provided lists of current students, there are an estimated 262 students in training. NLM has supported or is supporting 160 of those students still in training, or a little over 61 percent.

Post-doctoral Candidates

Six programs provided a list of alumni and 9 identified current postdoctoral students that met the established specifications. In total, the six programs that identified alumni have had approximately 117 graduates over the past 10 years. NLM funding supported 69 of trainees at those programs, or approximately 59 percent. Among current postdoctoral students, there are 92 trainees at the 9 programs, 47 of which have received or are receiving NLM funding.

Table 3: Number of Alumni and Current Trainees at Select Universities

	Number of Universities	Number of Trainees	Number of NLM-Funded Trainees	Percent NLM-Funded
Pre-doctoral Alumni	9[a]	289	174	60.2
Pre-doctoral Current	11[b]	262	160	61.1
Postdoctoral Alumni	6[c]	117	69	59.0
Postdoctoral Current	9[d]	92	47	51.1

a: Columbia University, Stanford University, University of California – Irvine, University of California – Los Angeles, University of Colorado, University of Virginia, University of Washington, University of Wisconsin-Madison, Yale University

b: Columbia University, Stanford University, University of California – Irvine, University of California – Los Angeles, University of Colorado, University of Missouri – Columbia, University of Pittsburgh, University of Virginia, University of Washington, Vanderbilt University, Yale University

c: Columbia University, Indiana University, Stanford University, University of California – Irvine, University of Colorado, University of Washington

d: Columbia University, Indiana University, Stanford University, University of California – Irvine, University of Colorado, University of Pittsburgh, University of Utah, University of Washington, Vanderbilt University

ASSESSING DEMAND: JOB ADVERTISEMENTS

The methodology used by Stephan and Black was adopted for this report. We examined job advertisements in the Back Matter section of *Science* magazine from 2001 to 2008. The purpose of the exercise was to examine if the number of jobs posted each year for individuals with degrees in biomedical informatics or related fields exceeded the number of graduates. It was also important to determine whether the jobs required skills that were refined in the biomedical informatics training programs. The goal of the NLM grants is to train scientists to do basic and applied research in the field, not to become hospital administrators or work as information technology workers. This exercise would determine whether the jobs available in the field allowed graduates of these training programs to conduct research.

Biomedical informatics is a difficult term to define. [2] This made it unwise to simply examine those advertisements seeking individuals with training in biomedical informatics. The search was expanded to include terms like computational biology, biostatistics, and informatics. Due to the interdisciplinary nature of the field, advertisements were included in the search from

---

[2] Biomedical informatics is an interdisciplinary field influenced by a variety of academic disciplines ranging from medicine, biology, computer science, and statistics. The interdisciplinary nature of the field makes it difficult to fully define. In "What is Biomedical Informatics?" published in the *Journal of Biomedical Informatics*, Doctors Elmer V. Bernstam, Jack W. Smith, and Todd R. Johnson addressed the obstacles faced when trying to develop a clear working definition of the field.

Bernstram et al. highlighted the deficiencies of previous biomedical informatics definitions in their paper. Past definitions concentrated on the technologies utilized by practitioners. The authors of the paper label these information technology-oriented definitions. By focusing on technology, there is a tendency to view computers and software as the solution to all problems and does not properly acknowledge work that does not involve computers. Role, task or domain-oriented definitions describe where informaticians are placed in an organization, making it difficult to consider tools developed for one purpose to be used in another domain. Concept-oriented definitions of biomedical informatics address the meaning of data, information and knowledge. These definitions are too esoteric to properly guide a scientific discipline (Bernstram, Elmer V. 2010).

The authors finally determine informatics is "the science of information" or "data with meaning." (The data can include anything from DNA sequences to electronic health records.) Drawing on that definition biomedical informatics is the science of information as applied to medicine and bio-medicine (Bernstram, Elmer V. 2010, p. 5). The definition is similar to the one adopted by the American Medical Informatics Association (AMIA), the professional association for this field. AMIA defines it as, "…the [study] and [pursuit] of the effective uses of biomedical data, information, and knowledge for scientific inquiry, problem solving and decision making" (American Medical Informatics Association 2011). Both definitions stress how data and information can be used to improve health outcomes.

the life sciences and computer sciences. Advertisements were rejected if the postings pertained to agriculture, veterinary medicine, neuroscience, or had no medical application.

These are lower-bound estimates. Each job was only counted once, even if it was advertised in multiple editions of *Science* magazine. Many of the postings advertised for multiple positions without giving an exact number of positions open. Those advertisements could only be counted once because the reviewers had no way of knowing how many positions companies were seeking to fill as far as ten years ago.

In addition to counting the advertisements seeking employees in the field, we also collected information on other characteristics. We categorized the advertisements by the type of institution hiring. Categories included colleges and universities, hospitals and other healthcare institutions, private industry, non-profit research organizations, and the federal government. If a hospital was affiliated with a school, it was counted as an academic institution. A few jobs remained uncategorized because little information was made available about the hiring institution.

Jobs were also categorized based on whether the individual was expected to work in the United States. Those jobs were labeled as "domestic." Jobs involving work outside of the United States were considered international. We also collected data on whether a doctorate degree was considered one of the prerequisites for the job and if it involved research.

Information was collected on 1,717 jobs from *Science* magazine. The number of jobs available in this field steadily declined between the years 2001 and 2008. In 2001, 428 jobs were advertised. This number decreased to 150 by 2008. On average, 214 jobs were posted each year.

Colleges and universities purchased the most job advertisements, accounting for 63.8 percent of all job advertisements posted during this time period. Nearly all of the jobs, 95.5

percent, involved research, while 87.4 percent of the jobs required applicants to have a doctorate degree. There were very few international jobs in the sample. Only 206 jobs involved work outside the United States.

Table 4: Science Magazine Job Advertisements

	2001	2002	2003	2004	2005	2006	2007	2008	Total
Total Jobs Posted	428	195	175	184	222	208	155	150	1,717
College/Universities	210	118	105	136	155	138	121	113	1,096
Hospital	2	1	2	1	7	3	0	2	18
Private Industry	162	54	43	16	26	26	13	16	356
Non-profit Research	15	8	1	9	7	20	7	1	68
Government	27	8	24	19	27	21	14	18	158
Not Categorized	12	6	0	3	0	0	0	0	21
Performs Research	404	192	175	169	214	200	149	136	1639
Requests PhD	412	195	175	146	182	162	121	107	1,500
International Jobs	56	26	19	31	27	0	18	29	206

The decline in the number of jobs advertised in *Science* magazine may not be due to the job market shrinking for individuals with advanced degrees in biomedical informatics. Two other theories could explain this decline. Job searches for most individuals could have shifted online, causing employers to feel it is less necessary for them to advertise in print media. The magazine even created its own online web page devoted to job advertisements. The job search could have been conducted there. The second theory is that the cost of placing an advertisement in a print magazine could have been too expensive. A print advertisement in *Science* magazine costs $520 for a ten-line advertisement, while an eight-week online advertisement only costs $425. Employers could have discovered cheaper, more cost-effective methods of informing applicants of job opportunities. The decrease could also be attributed to the declining popularity of the magazine.

To test these theories, an online source, Bioinformatics.org, underwent a similar examination. Archives of job posting were only available as far back as December 2006, but this was a better time frame than any other online job site considered. The years 2007, 2008, 2009, and 2010 were considered. Almost 1,000 job advertisements were found meeting the criteria outlined above. Unlike in *Science* magazine, the research team observed an increase in the number of job postings on Bioinformatics.org. In 2007, 219 advertisements were posted. This number climbed to 312 in 2010. This source indicates that the job market for individuals with advanced degrees in biomedical informatics is robust, and grew between 2007 and 2010.

Table 5: Bioinformatics.org Job Postings

	2007	2008	2009	2010	Total
Total Jobs Posted	219	231	234	312	996
College/Universities	79	82	104	116	381
Hospital	25	17	16	13	71
Private Industry	75	86	55	122	338
Non-profit Research	10	15	34	45	104
Government	24	24	22	12	82
Not Categorized	6	7	2	4	19
Performs Research	143	148	191	263	745
Requests PhD	120	125	154	186	585
International Jobs	65	53	46	45	209

There was a greater amount of diversity in the number of job postings at this source. While most of the jobs advertised in *Science* magazine were for work at academic institutions, only 38.3 percent of the jobs were at academic institutions on Bioinformatics.org. Private corporations advertised for 33.9 percent of the jobs on Bioinformatics.org. More international organizations and academic institutions advertised on the site. Nearly 21 percent of the jobs were for organizations in foreign countries, such as Singapore and India. Approximately 75 percent of the jobs involved some research and 58.7 percent of the jobs considered a PhD to be one of the requirements for applying for the position. Despite the presence of more diversity in the types of organizations who sought employees on this website, most of the jobs were still for research-oriented positions and required PhDs, which is similar to what was observed in the *Science* magazine job advertisements.

When we consider the total number of job postings for both the print and online sources we examined, we find that there were a total of 2,713 relevant advertisements over a 10 year period. For the reasons explained in the methodology section, this number is a lower bound estimate, meaning that the actual number of jobs available is higher than 2,713. Our estimate for the supply of trainees over the same time frame is 929 individuals. As explained in the supply section, this number is potentially an overestimate because of the assumption that programs have graduated similar numbers of trainees at a similar rate over 10 years. These two estimates reveal a substantially higher number of job advertisements than trainees, indicating a strong demand.

Survey Results

Evidence gathered from lists of program participants and alumni, as well as the job advertisement data can provide one aspect of the labor market for PhD biomedical informatics. What this approach lacks is the first hand experience that trainees actually face when they complete their degrees. We developed a survey in order to fill in this void. The directors of the eighteen NLM-funded programs were contacted via email with a letter explaining the project and asking for their help to distribute the survey. Twelve of the directors agreed and distributed the survey to their alumni through a provided hyperlink. There was a fair response to the survey: 116 participants completed the survey, 106 of whom had earned an advanced degree. Approximately 93 to 96 of those participants who earned a degree answered each question. Respondents were not required to answer every question, leading to this disparity in responses per question.

Who answered the survey?

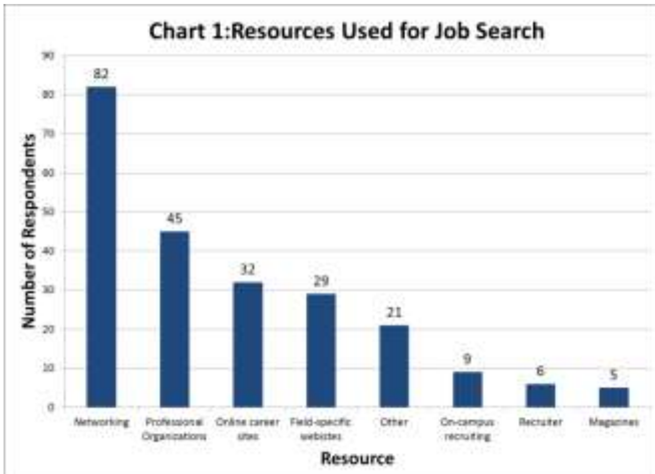
As our major charge revolved around the PhD labor market, we marked the survey towards those individuals who had earned doctorate degrees from biomedical informatics training programs. In anticipation of non-PhD seeking individuals participating, the first question asked the respondent to identify their relationship to the biomedical informatics program they attended. Approximately 62 percent of respondents were pre-doctoral candidates who earned a PhD, while 11 percent had earned a PhD, MD, or equivalent and received an additional PhD. Nine percent of respondents did not earn a degree, while there was 9 percent of respondents who earned a master's after having earned a PhD, and an additional 9 percent who entered pre-doctoral and earned only a master's degree.

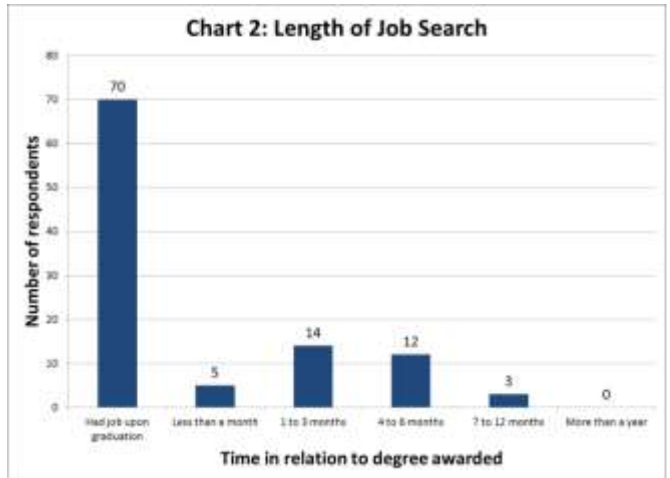
The sample was heavily male, with only 34 percent of respondents being female. The average age when respondents received their degree was 32 years. There do not appear to be large outliers influencing this number, as the median age is 31 years old. While there were some respondents who had received their degree prior to the year 2000, the majority of participants (72.6 percent) received their degree between 2001 and 2011. This will allow crucial insight into the experience recent graduates have had in the labor market. For those who entered both a year that they began their training and a year that they ended their training, the average length of training was a little over 4.5 years.

As a way to identify NLM-funded trainees, the survey asked respondents to enter their name. The research group was able to positively identify 62 participants as individuals who had received training support from NLM over the past ten years. Of the remaining participants, 19 chose to enter "anonymous" or not to enter a name, 4 entered only a first name and the remaining 31 were not NLM-funded trainees during the past ten years. It is entirely possible that the anonymous or semi-anonymous participants could have been funding recipients. Since they could not be positively identified in either way, when there are discussions between NLM funded and non-NLM funded trainees, these individuals will be excluded. Similarly, due to data limitations, some of those who earned a degree prior to 2001 may have been funding recipients without the research team having the ability to check. These individuals will not be included in either the funded or not funded groups.

The Job Search

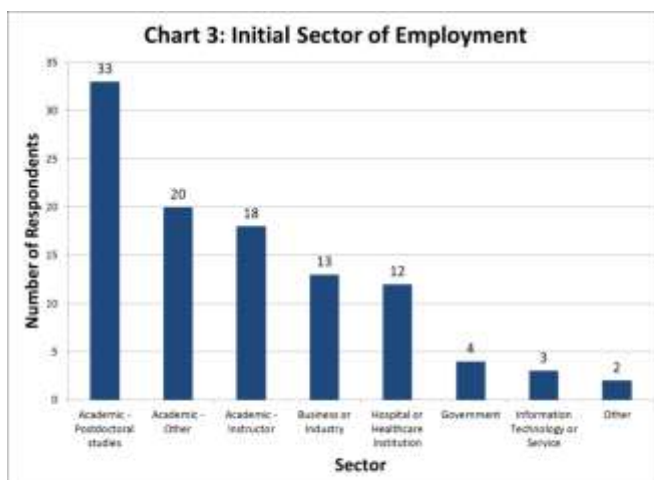
When asked what resources they used in their job search, the most popular response was networking, with 82 respondents. Among the 21 who selected "Other," six specified responses that could also be considered networking such as conferences, colleagues, word of mouth, PhD advisor and personal contacts. The next most popular responses were professional organizations (45), online career sites (32), and field-specific websites (29). On-campus recruiting (9), recruiters (6), and magazines (5) appear to be lesser used resources for biomedical informaticians. These results further enforce the notion that the research group's analysis of job postings are lower bound estimates and the downward trend in print advertisements, since only 5 respondents selected magazines as a job hunting tool. Job opportunities are probably higher than our estimates due to the large number of respondents who selected networking as a job hunting tool.





Long job searches do not appear to affect this field. The majority of respondents, 70, had a job upon graduation, while all but 3 had a job within 6 months and none had to wait more than a year. This is very good news to trainees as it would appear to confirm anecdotal evidence that these individuals are highly demanded.

Academic jobs are the most prevalent first employment sector for these respondents, with Academic-Postdoctoral studies being the most popular (33), followed by Academic-Other (20) and Academic-Instructor (18). The Academic-Other category asked the respondents to clarify, with 7 selecting professor or assistant professor and 8 specifying some type of researcher position. Very few entered government service (4) or the information technology sector (3). Thirteen began their career in business or industry and 12 started in a hospital or healthcare institution.



One of the main goals of the NLM training grant program is that the trainees find research based jobs after they receive their degrees. Reports from each university give the current job title and employer for past funding recipients, if available. Job titles do not adequately convey actual job responsibilities. A question on the survey asked respondents to provide a description of their major responsibilities at their initial job. From the research team's tabulations of these responses, 63 respondents indicated research as a primary responsibility of their first job. Twenty cited teaching, 16 developed informatics tools, 17 had been involved in health information technology, and 9 indicated applying for grants as a major responsibility. Only 3 entered responses that did not appear to be job descriptions or responsibilities associated with biomedical informatics.

Job History

Fifty-five respondents are currently with their first employer. This is perhaps reflective of the portion of the sample that has only received their degree within the last 10 years. Aside from these individuals, the initial job does not appear to be viewed as a long-term position, with 35 respondents staying with their first employer for 1 to 5 years. Only 8 respondents stayed with their first employer for more than 5 years, and 5 stayed for less than a year.

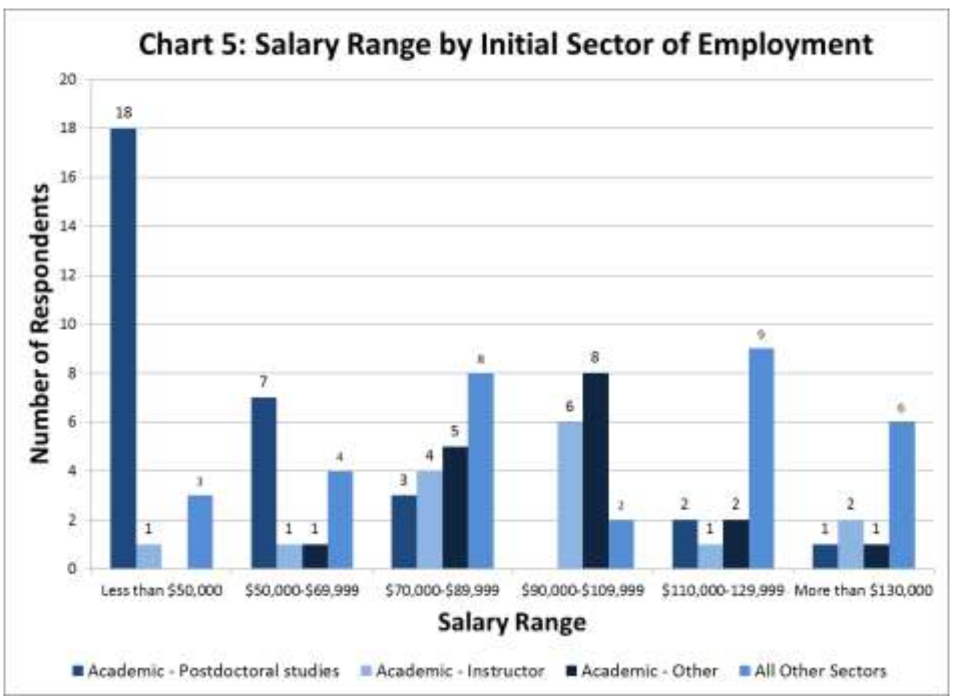
Among those still with their initial employer, there appears to be some degree of intra-mobility. Fourteen have been in their current position for less than a year, while 32 have been in their position for 1 to 5 years. Eight have had the same position for more than 5 years.

The combination of the first employment sector and length of time with the first employer can provide some insight into questions currently being addressed by the Chalkley report. One issue being considered in that report is that PhD researchers are being held in postdoctoral positions. First, we restricted our sample to only those who selected Academic-Postdoctoral studies as their initial employment sector. Among these 33 individuals, only one stated that they were with that employer for more than 5 years. Seventeen respondents stayed in this sector for 1 to 5 years, but 14 of those individuals graduated between 2007 and 2009, indicating that they were there for some time less than 4 years. There are 12 individuals who are currently still with their initial employer. Seven are recent 2010 or 2011 graduates. Four graduated between 2005 and 2009 and one graduated pre-2000. Breaking this subgroup down further, there are 6 that have been in their current position for less than a year. Five of those respondents graduated in 2010 or 2011, which means that at the time of the survey they may not have yet been out of school for more than a year. The remaining 6 have all been in their current position for 1 to 5 years and all except for one graduated before 2010. These statistics do not appear to imply that long postdoctoral appointments seriously afflict the biomedical informatics labor market.

Next, those respondents who had ceased employment at their first job were asked how many employers they had had since receiving their degree. The most common response was 2 employers (27), followed by 1 employer (10). As this question was only asked to those who did not select "currently with first employer," the response of one employer could indicate two things. First, these respondents may have ceased employment with their first employer and are currently between jobs. The second possibility is that this is an instance of respondent error. Aside from these two responses, 6 respondents had 3 employers while there were each one instance of 4, 6, and 7 employers. Among the 9 respondents who had 3 or more employers, all except for 1 had graduated before 2004. Four graduated during the 1990s. Five of these respondents initially worked in the business or industry, hospital or healthcare, or information technology sectors.

Respondents were also asked to indicate the salary range of their first job. Twenty-two indicated making less than $50,000 at their first job, while 10 made more than $130,000. The rest fell fairly evenly between the remaining four brackets. There are some facts that can be gained by combining these answers with answers from previous questions. First, one economic theory states that in a labor market that appears to be as strong as this one, some individuals may take longer to find a job in an effort to earn a higher salary. The research team broke down the salary ranges by the length of the job search and found that this does not appear to be the case. Higher salaries were not clustered in the later time periods, though this analysis may be impeded by the small number of individuals with extended job searches.

Second, salaries might also be able to tell information about future funding decisions by NLM. When the costs of training are subsidized, it encourages more individuals to enter the field. If, however, individuals are able to earn high salaries once they end their training and thus pay back the costs of their education, there is less of a need for the field to be subsidized.



The data here presents a mixed picture, with a wide variation in initial salaries. Some of the variation can be explained by the initial job sector. Of the 22 individuals making less than

$50,000, 18 took academic post-doctoral positions. The other initial employment sectors were not clustered as strongly as the academic post-doctoral sector, but this is also tempered by the fact that some sectors had a very small number of respondents. Educational path also runs into a similar problem. The variation of salaries remains fairly uniform, but this may be due to the lower number of respondents for the categories other than pre-doctoral candidates earning a PhD.

The final question most respondents saw asked them whether they had experienced a gap in employment that lasted longer than 6 months. Only 5 respondents indicated that they had, and when asked why, 1 indicated personal reasons, another indicated being laid off, the third said a lack of job opportunities, and another respondent selected "other." The final one did not specify a reason.

Another question of interest about the field was how trainees disseminate their research. By far, the most popular method is through published articles with 93 respondents. Informatics tools came in second with 42 responses. Databases of scientific data and patents had a smaller percentage of respondents, with 22 and 17, respectively. Twenty-two respondents selected "other" and their responses were generally conferences or other presentations.

We attempted to conduct several comparison group analyses. First, we considered non-NLM-funded trainees and NLM-funded trainees. There were only 10 individuals who could be positively identified as non-NLM-funded trainees, which limited potential comparisons. These respondents did not exhibit any meaningful differences or instances of clustering around specific answers.

Similarly, when respondents were separated by the four potential educational paths, the small sample sizes for everything except for "pre-doctoral candidate who earned a PhD" limit our analysis. The strongest instance of clustering among the other three educational paths occurred with pre-doctoral candidates who earned only a master's degree. Out of those 11 respondents, 6 found initial employment in a hospital or healthcare institution. These individuals represent half of those respondents who selected that sector. Aside from this point, other responses followed the patterns of the full sample or were spread out evenly.

Finally, in regards to the individuals who indicated that they did not earn a degree, only one indicated that he/she did not finish the degree due to limited financial resources. Of the other nine, six indicated that they were non-degree seeking students. One indicated that he/she began work on a start-up company and another explained a difficulty with the degree granting institution.

Implications

There are several important factors to take away from this analysis. First, individuals who receive advanced training in biomedical informatics quickly find employment and there are few instances of long term unemployment. Second, when considering the job advertisement and trainee data, even if the number of alumni calculated in the analysis were doubled (which assumes that the programs all have similar numbers of alumni), there would still be fewer total alumni over the past 10 years than an aggregation of job advertisements over that time frame. Third, these trainees do not appear to be being held in lengthy post-doctoral research positions; they readily enter the labor force primarily as academic professors or instructors, as well as industry, healthcare, and government. Given these factors, there is no reason to limit the number of trainees.

This analysis, particularly the survey portion, is limited by the sample size and make-up. Based on the data gathered and the trainees' experiences in the labor market, demand appears to be strong and there does not appear to be a surplus of trainees. These facts lend credence to the National Library of Medicine's goal of increasing the number of biomedical informaticians.

Recommendations

The survey and analysis of job posting provided a snapshot of the labor market for individuals with advanced degrees in biomedical informatics. More details about the labor market need to be known, however, to aid NLM and other governmental organizations in deciding future funding amounts for these training programs. Outlined below are suggestions for further research.

Track the total number of biomedical informatics graduates.

NLM should track the number of graduates receiving master's and doctorate degrees in biomedical informatics each year from programs that offer training in at least one of the four training areas. Currently, NLM only knows the number of students its training grants supports. Several schools list their alumni on their program websites. This is not a common or thorough enough practice to produce a comprehensive list of graduates of biomedical informatics programs.

Even within the academic intuitions that receive NLM grants, there are students not supported by grants. For a full understanding of the supply of individuals with advanced degrees with biomedical informatics, everyone should be counted. NLM will then have a better idea of the proportion of trainees that its grants support. In the future and assuming a constant level of funding, if NLM observes an increase in the total number of trainees while the proportion of trainees it supports decreases, this could indicate that the field is maturing and resources could be directed elsewhere.

Distribute a survey to employers in the field.

Past labor market studies focused on graduates with advanced degrees in biomedical informatics. Very little is known about the employers' experiences and preferences. A survey could be distributed to major employers of individuals with advanced degrees in biomedical informatics. These employers include major research universities and large pharmaceutical companies, such as Eli Lilly.

The survey would be used to learn if employers face any challenges while recruiting informaticians. More details about the job market can be uncovered including: which resources are used to recruit informaticians, how long does it take to fill positions, whether they are seeking individuals with doctorate degrees, and whether these companies have to recruit abroad to fill positions. If the institutions surveyed say that it is difficult to find individuals with training in biomedical informatics, then it would indicate that the supply of individuals with advanced degrees in biomedical informatics should increase.

More importantly, the survey would help determine why colleges and universities appear to be the most popular employers of biomedical informaticians. In Stephan and Black's earlier labor market research for the biomedical informatics field, they discovered private organizations and companies advertised for the most positions in *Science* magazine in 1996 and 1997. By the year 2002, most of the advertisements were for positions at academic institutions. A survey to the major employers in the field would seek to determine whether there is still a demand for the services of informaticians in the non-academic private sector.

Determine how many full-time positions are funded by NLM and other governmental organizations.

To fully understand what is driving demand for biomedical informaticians, it must be understood how many jobs are being funded by public funds. Public funding plays a significant role in scientific research. This information can be used to better understand the landscape and answer the following questions: Is it necessary for the federal government to support the field before it matures and becomes lucrative? Is there a real private sector demand for this skill set? Is the federal government crowding out private investment? These questions will not be answered without examining how many positions are funded by the National Institutes of Health and other federal agencies.

REFERENCES

American Medical Informatics Association. (2011). *Academic forum members*. Retrieved September 19, 2011, from http://www.amia.org/programs/academic-forum/members

American Medical Informatics Association. (2011). *Biomedical informatics core competencies*. Retrieved September 19, 2011, from http://www.amia.org/biomedical-informatics-core-competencies

Basil, M. D., & Basil, D. Z. (2006). The marketing market: A study of PhD supply,demand, hiring institutions, and job candidates. *Journal of Business Research, 59*(4), 516-523. doi:10.1016/j.jbusres.2005.08.005

Bernstram, E. V., Smith, J. W., & Johnson, T. R. (2010). What is biomedical informatics? *Journal of Biomedical Informatics, 43*(1), November 28, 2011-14. doi:10.1016/j.jbi.2009.08.006

Black, G. C., & Stephan, P. S. (1999). Bioinformatics: Does the U.S. system lead to missed opportunities in emerging fields? A case study. *Science and Public Policy, 26*(6), 1-30.

Black, G. C., & Stephan, P. S. (1999). *Hiring patterns experienced by students enrolled in Bioinformatics/Computational biology programs*. Alfred P. Sloan Foundatio:

Black, G. C., & Stephan, P. S. (2005). Bioinformatics training programs are hot but the labor market is not. *Biochemistry and Molecular Biology Education, 33*(1), 58-62. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/bmb.2005.494033012428/pdf

Center for Comprehensive Informatics, Emory University. (2011). *About the center*. Retrieved November 24, 2011, from http://cci.emory.edu/cms/about/index.html

Grimes, M. F., & Grimes, P. W. (2008). The academic librarian labor market and the role of the master of library science degree: 1975 through 2005. *The Journal of Academic Librarians, 34*(4), 332-339. doi:10.1016/j.acalib.2008.05.023

Hersh, W. (2006). Who are the informaticians? *Journal of the American Medical Informatics Association, 13*(2), 166-170.

Hersh, W. (2008, Health and biomedical informatics: Opportunities and challenges for a twenty-first century profession and its education. *IMIA Yearbook of Medical Informatics,* 157-164.

Knaup P. (2003). Medical informatics specialists: What are their job profiles? results of a study on the first 1024 medical informatics graduates of the universities of heidelberg and heilbronn. *Methods of Information in Medicine, 42*(5), 578-587.

Patton, G. A., & Gardner, R. M. (1999). Medical informatics education: The university of utah experience. *Journal of the American Medical Informatics Association, 6*(6), 457-465.

The National Library of Medicine. *NLM's university-based biomedical informatics research training programs*. Retrieved September 3, 2011, from http://www.nlm.nih.gov/ep/GrantTrainInstitute.html

The National Library of Medicine. (2011). *Fiscal year 2011 awards*. Retrieved September 3, 2011, from http://www.nlm.nih.gov/ep/Awards2011.html#Uni

The National Library of Medicine. (2011). *NLM institutional training grants for research training in biomedical informatics (T-15).* Retrieved September 5, 2011, from http://grants.nih.gov/grants/guide/rfa-files/RFA-LM-11-001.html

Appendix A[3]

National Library of Medicine Funding for Fiscal Years 2001-2011

Institution	Total	FY 2011	FY 2010	FY 2009	FY 2008	FY 2007
Columbia University Health Sciences	$17,444,065	$1,101,038	$922,661	$1,449,520	$1,407,104	$1,352,260
Harvard University (Medical School)	$20,479,392	$1,414,429	$2,010,461	$1,895,746	$1,510,011	$3,088,288
Indiana Univ-Purdue Univ at Indianapolis	$4,505,433	$455,020	$360,437	$536,961	$414,860	$512,521
Johns Hopkins University	$5,265,129	$93,050	$358,231	$598,896	$403,498	$543,922
Oregon Health and Science University	$7,282,316	$603,939	$844,608	$762,022	$702,910	$682,856
Rice University	$8,319,208	$674,587	$507,290	$1,878,390	$615,216	$877,130
Stanford University	$11,065,291	$939,534	$1,107,126	$1,213,351	$967,500	$2,409,080
University of California Irvine	$9,151,297	$694,958	$1,058,574	$858,556	$884,083	$1,008,201
University of California Los Angeles	$5,819,875	$525,750	$310,541	$1,077,670	$518,696	$525,880
University Of Colorado Denver	$3,066,191	$406,924	$726,019	$962,875	$497,807	$472,566
University of Missouri-Columbia	$5,854,413	$151,980	$597,160	$941,891	$145,176	$813,651
University of Pittsburgh at Pittsburgh	$13,487,551	$800,312	$694,902	$2,081,185	$1,045,638	$1,195,518
University of Utah	$10,073,544	$882,483	$974,438	$1,239,467	$1,044,704	$1,060,610
University of Virginia	$3,917,916	$459,424	$766,894	$970,112	$689,434	$1,032,052
University of Washington	$9,459,970	$960,328	$750,264	$1,147,919	$1,272,285	$1,137,852
University of Wisconsin Madison	$11,178,566	$975,085	$954,573	$1,040,693	$1,090,904	$1,743,582
Vanderbilt University	$8,954,077	$663,943	$663,216	$1,511,873	$635,108	$1,813,143
Yale University	$10,899,946	$961,444	$1,089,873	$1,189,449	$938,744	$907,779
Total	$166,224,180	$12,764,228	$14,697,268	$21,356,576	$14,783,678	$21,176,891

[3] Information accessible at: http://report.nih.gov/index.aspx

Institution	FY 2006	FY 2005	FY 2004	FY 2003	FY 2002	FY 2001
Columbia University Health Sciences	$1,216,329	$2,742,168	$2,206,094	$2,546,081	$1,721,842	$778,968
Harvard University (Medical School)	$2,035,431	$2,313,652	$1,682,476	$1,618,207	$1,480,547	$1,430,144
Indiana Univ-Purdue Univ at Indianapolis	$422,452	$371,328	$399,681	$243,476	$389,673	$399,024
Johns Hopkins University	$807,569	$820,960	$632,731	$667,340	$338,932	
Oregon Health and Science University	$757,696	$644,816	$744,401	$657,180	$454,206	$427,682
Rice University	$810,595	$714,297	$341,383	$768,189	$546,760	$585,371
Stanford University	$141,801	$401,033	$961,480	$975,957	$899,770	$1,048,659
University of California Irvine	$950,305	$1,195,553	$1,075,541	$869,165	$556,361	
University of California Los Angeles	$575,908	$494,862	$843,948	$774,264	$172,356	
University Of Colorado Denver						
University of Missouri-Columbia	$637,442	$381,484	$626,178	$556,343	$365,282	$637,826
University of Pittsburgh at Pittsburgh	$1,292,900	$1,384,140	$1,133,556	$1,550,250	$1,187,402	$1,121,748
University of Utah	$796,892	$987,435	$724,428	$989,555	$717,257	$656,275
University of Virginia						
University of Washington	$1,013,437	$1,113,363	$931,422	$759,172	$373,928	
University of Wisconsin Madison	$1,087,860	$1,108,085	$1,088,720	$1,098,062	$991,002	
Vanderbilt University	$660,515	$674,904	$946,630	$776,591	$608,154	
Yale University	$1,256,605	$1,043,144	$966,927	$1,188,817	$758,564	$598,600
Total	$14,463,737	$16,391,224	$15,305,596	$16,038,649	$11,562,036	$7,684,297

Appendix B

Survey Copy

Welcome Message: A team of researchers at The College of William and Mary's Thomas Jefferson Program in Public Policy is investigating the labor market for PhD recipients in biomedical informatics for the National Library of Medicine. Your participation is key in helping to better understand the demand for research-trained biomedical informaticians.

The Thomas Jefferson Program in Public Policy understands that your privacy is important to you. Your information will be kept confidential. The results of the survey will only be used for scholarly purposes.

Thank you for your participation. We look forward to reading your responses!

Question 1: Name

Question 2: Did you receive a PhD in biomedical informatics or a related informatics field?

- Yes, I was a predoctoral candidate who received a PhD.
- Yes, I entered the graduate program after having earned a PhD, MD, or equivalent and earned another PHD.
- No, I entered the graduate program after receiving a PHD and earned another master's degree.
- No, I was a predoctoral candidate who received only a master's degree.
- No, I did not earn a degree. (GO TO QUESTIONS 21 through 24)

Question 3: Please enter the age when you received your degree in biomedical informatics or a related field. (Respondent enters.)

Question 4: Gender

- Male

- Female

Question 5: At which academic institution did you receive your degree in biomedical informatics or a related informatics field? (Respondent enters.)

Question 6: Please enter the month and year you entered your graduate program in biomedical informatics. (Month and year drop down boxes.)

Question 7: Please enter the month and year you were awarded your degree in biomedical informatics or a related field. (Month and year drop down boxes.)

Question 8: What resources did you use in your job search? (Please check all that apply.)

- Magazines
- Online career sites
- Professional organizations (mailings or websites)
- Websites targeted for individuals in my field
- Networking
- On-campus recruiting
- Recruiter
- Other (Please specify)

Question 9: In thinking about the first job you had following your degree, about how long did your job search take?

- Had job upon graduation
- Less than a month following graduation
- 1 to 3 months
- 4 to 6 months
- 7 to 12 months
- More than a year

Question 10: How would you characterize the employment sector of your first job?

- Academic – Instructor
- Academic – Postdoctoral studies
- Academic – Other (Please specify)
- Business or industry
- Hospital or healthcare institution
- Government
- Information technology or service
- Other (Please specify)

Question 11: Briefly describe the major responsibilities at your first job. (Respondent enters.)

Question 12: How long did you stay with your first employer?

- Currently with first employer (GO TO QUESTION 13)
- Less than a year (GO TO QUESTION 14)
- 1 to 5 years (GO TO QUESTION 14)
- More than 5 years (GO TO QUESTION 14)

Question 13: If you are currently with your first employer, how long have you been in your current position?

- Less than a year (GO TO QUESTION 16)
- 1 to 5 years (GO TO QUESTION 16)
- More than 5 years (GO TO QUESTION 16)

Question 14: How many employers have you had in biomedical informatics or a related field since graduation? (Drop down box.)

Question 15: If you have changed employers since your initial job, how would you characterize the employment sectors in which you have worked? (Please select all that apply.)

- Academic – Instructor
- Academic – Postdoctoral studies
- Academic – Other (Please specify)
- Business or industry
- Hospital or healthcare institution
- Government
- Information technology or service
- Other (Please specify)

Question 16: How do you disseminate the results of any research you conduct? (Please select all that apply.)

- Published articles
- Informatics tools
- Database of scientific data
- Patents
- Other (Please specify.)
- Not applicable

Question 17: Please select the income range that best describe the annual salary of your first job. (Drop down box of income brackets)

- Less than $30,000
- $30,000-$49,999
- $50,000-$69,999
- $70,000-$89,999
- $90,000-$109,999
- $110,000-129,999
- $130,000-$149,999

- More than $150,000

Question 18: Have you experience any gaps in employment lasting six months or longer since your first job?

- Yes (GO TO QUESTION 19)
- No (This ends the survey)

Question 19: Approximately how long did your longest period of joblessness last? (Example: Six months, Two years, etc.) (GO TO QUESTION 20)

Question 20: Please select the primary cause for your longest period of joblessness.

- Personal/family reasons (This ends the survey)
- Lack of job opportunities (This ends the survey)
- Employer ceased operation (This ends the survey)
- Laid off (This ends the survey)
- Other (Please specify.) (This ends the survey)

Question 21: Why did you not complete you degree in biomedical informatics or a related informatics field?

- Limited financial resources
- Personal reasons
- Decided to pursue another field (Please specify the field in the box below.)
- Other (Briefly describe below.)

Question 22: Did you seek a job after you ended your training?

- No
- Yes

Question 23: How long did your job search take?

- Received job offer while in training

- Less than a month after training ended
- 1 to 3 months
- 4 to 6 months
- 7 to 12 months
- More than a year

Question 24: How would you characterize the employment sector of your job?

- Academic – Instructor
- Academic – Postdoctoral studies
- Academic – Other (Please specify)
- Business or industry
- Hospital or healthcare institution
- Government
- Information technology or service
- Other (Please specify)

(Question 24 is the final question if this response thread was followed.)