

Vedana Vaidhyanathan Why Can't It All Be On the Web?: The Information Needs of Biomedical Informatics Scientists. A Master's paper for the M.S. in L.S. Degree. July, 2008. 42 Pages. Advisor: Barbara Wildemuth.

Bioinformatics researchers are a relatively new breed of scientist, emerging from many disciplines. In order for libraries to meet the needs of these researchers, a clear assessment of how they use the library and how they would like to use it is essential. Through in-person interviews with graduate students, post-doctoral students, researchers and medical doctors, this study examines bioinformatics researchers' use of the library and what would constitute their ideal library. It asks them how they currently use the library, and how they would change the library to better serve their research needs. The findings reveal that most of the researchers want more access to online sources and they want an easier way to access these sources than is currently provided. It is recommended that further studies concentrate on how libraries can publicize their current resources to target the bioinformatics research population.

Headings:

Bioinformatics

Information Needs

Information Assessment

Library Assessment

Biomedical Informatics

Bioinformationist

“WHY CAN’T IT ALL BE ON THE WEB?”: THE INFORMATION NEEDS OF
BIOMEDICAL INFORMATICS SCIENTISTS

Vedana Vaidhyanathan

A Master’s paper submitted to the faculty
of the School of Information and Library Science
of the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements
for the degree of Master of Science in
Library Science.

Chapel Hill, North Carolina
July 2008

Approved by:

Barbara Wildemuth

ACKNOWLEDGEMENTS

To Barrie Hayes, K.T. Vaughan and all the librarians at the UNC Health Sciences Library, thank you for giving me the opportunity to explore this issue. Without your support and guidance I would not have gotten to where I am today.

To my advisor Dr. Barbara Wildemuth, thank you for giving me your guidance and expertise as I navigated through this process. I will never forget your support.

To my parents, my father Vishnampet Sivaramakrishnan Vaidhyanathan and my mother Virginia Ann Vaidhyanathan, thank you for always believing in me, never giving up on me and supporting me through late-night phone calls, reassuring pep talks and many hugs. I could not have made it through graduate school without you.

To Abbey Thompson, for making everything happen.

And to Jaya Francesca Vaidhyanathan, for inspiring me purely by existing.

TABLE OF CONTENTS

List of Tables.	3
List of Figures	4
Introduction	5
Literature Review.	7
Methods.	14
Identifying and recruiting the study participants.	15
Conducting the interviews.	17
Analyzing the interview data.	18
Results and Discussion	21
Limitations of the study.	33
Conclusion.	34
Bibliography	37
Appendix	39

LIST OF TABLES

Table 1: Experience at UNC.	20
---	----

LIST OF FIGURES

Figure 1: Primary discipline.	21
Figure 2: Status of participants.	22
Figure 3: Departmental appointments of study participants.	23
Figure 4: Physical library last visited.	24

Introduction

Bioinformatics is a field of study that is relatively new. It rose to national and international prominence when the human genome was sequenced using tools that had not previously been in common practice prior to its sequencing. The Oxford English Dictionary Online (<http://dictionary.oed.com/>) defines bioinformatics as “the science of information and information flow in biological systems, esp. of the use of computational methods in genetics and genomics.” This is a very general definition, and encompasses many different programs and disciplines. The University of North Carolina at Chapel Hill is conducting interdisciplinary research in bioinformatics and participants in the School of Information and Library Science have provided their own definition:

Biomedical informatics is the application of information science and technology to the life sciences (biochemistry, biology, genetics, genomics, proteomics, etc.). Biomedical informatics has played a substantial role in large scale research initiatives such as the Human Genome Project. (SILS Bioinformatics Journal Club website, <http://bioivlab.ils.unc.edu/JoomlaBioinfo/>)

At the University of North Carolina, researchers in bioinformatics have been identified by the Health Sciences Library in the disparate fields of Computer Science, Biology, Biochemistry, Biophysics, Information Science, Pharmacy, Public Health, Botany, Dentistry and Medicine. There may be additional researchers in other departments that have yet to be identified.

Many health sciences and medical libraries have begun supporting research in bioinformatics. In some research universities this has been accomplished by employing

scientists with PhDs in molecular biology and subsequently training them to be librarians; other programs have chosen to take a complementary route, employing librarians and training them in the biomedical sciences.

In an attempt to understand what services libraries can provide to the bioinformatics community, libraries have conducted surveys questioning whether bioinformatics professionals are aware of the resources in the library and if they would like training in these resources. The Health Sciences Library at the University of North Carolina has conducted informal surveys to understand the bioinformatics community's awareness of the available resources, but the community is so disparate that it is hard to get a good sense of their needs based on the existing survey results. Because these scientists are dispersed among different departments, it is difficult to answer the question, "How do bioinformatics researchers use the University of North Carolina's libraries?" In order to answer this question, the current study used in-depth interviews. Such interviews could display general trends, but could also help us learn about more targeted issues for future growth and expansion of services. This study was conceived through conversations between the Bioinformatics Librarian, the Pharmacy Librarian and the Bioinformatics Fellow at the Health Sciences Library.

The purpose of this study, broadly defined, is to find out if bioinformatics researchers use the library system at UNC and, more specifically, the Health Sciences Library (HSL). The quality of their experiences at the HSL was examined, as well as the researchers' desires for present and future library resources and services. If they do use the library system, in which way(s) do they use it? Do they use it mainly through websites, or do they physically walk in and ask the librarian questions? This study also

explored what the researchers see as the primary functions of their library. Recent library scholarship has often focused on how patrons want access to information electronically, and no longer think print is important, expounding on the virtues of the library as the “third place.” Do scientists in the bioinformatics community also find these issues to be important?

From this study, the library hoped to find out more about the information needs of bioinformatics researchers and how they conduct research. The study also hopes to better identify the population of scientists who consider themselves bioinformatics researchers. The term bioinformatics is so ambiguous, it takes self-identification in order to know who, exactly, is in the population. The field of bioinformatics is interdisciplinary and cross-disciplinary so it is often difficult to establish a population without self-identification. Interdisciplinary means that the research attacks one subject from many different angles; for example bioinformatics could be considered interdisciplinary because it is used in biology, chemistry and many other disciplines. It is cross-disciplinary because it can be studied across many disciplines without there being a primary discipline attached to it. This study is one step in the process of identifying the information needs of the bioinformatics population, in order to better serve them.

Literature Review

How do bioinformatics scientists currently use and desire to use the library? That is the question this study hopes to answer. Are they aware of the different subscriptions, titles and databases the library possesses? Can they find the information they need through the library or must they use outside sources? Are they even aware that the library

provides sources for them? Before answering these questions within the UNC context, it is important to survey the prior literature.

Libraries have found they need to tailor their services in order to keep up with the challenges of the changing information environment. This literature review consists mainly of articles reporting on surveys executed to try to find out how to better serve the information needs of bioinformatics scientists and articles about conducting surveys on the information needs of other scientists. There are three main themes that permeate the previous literature: collaboration, training of librarians in bioinformatics, and the development of bioinformatics as a new field in which librarians can make a name for themselves.

Bioinformatics is a field in which librarians can expand their traditional roles and experiment with new roles in order to bring about change and collaboration between researchers in the field. “Medical libraries, bioinformatics, and networked information: a coming convergence?,” by Lynch (1999), is a research article, but it is not an empirical research article. It is an article about the future of medical libraries and how bioinformatics and networked information are coming together. It predicts a need for more instruction in science for medical librarians and a greater symbiosis between the medical library and the bioinformatics scientists. It is the touchstone of the research into how medical libraries can reach out to bioinformatics scientists.

Similarly, the article, “Genetic Information Resources: A New Field for Medical Librarians,” (Norman, 1999) talked about the future of bioinformatics and how there is a place for librarians in the post Human Genome Project world. This is important since it

expresses the hope that librarians will be at the forefront of the convergence of disciplines in bioinformatics.

The article, “Bioinformatics Opportunities for Health Sciences Librarians and Information Professionals,” by Helms et al. (2004), articulated the same idea. It talked of how the library can be at the forefront of bioinformatics research. It also articulated the ideas that come up in the articles with other themes: collaboration and training of librarians to be in bioinformatics.

Collaboration comes in many forms. It is always a partnership between the library and the bioinformatics programs on campus. It usually, though not necessarily, encompasses some training. In some cases collaboration is through providing training and workshops, in other cases it is through one-on-one consultations with faculty led by a specially created librarian, usually with a science or medical background who understands the researchers’ needs. Chattopadhyay et al.’s (2006) article entitled, “Design and Implementation of a Library-based Information Service in Molecular Biology and Genetics at the University of Pittsburgh,” is about one of those programs. In the article, which was written by a scientist who was hired into the University of Pittsburgh’s library, is an analysis of the workshops the library held and the surveys they conducted to get user feedback about those workshops. The article discussed the workshops and trends for the future, and analyzed their web portal. Unfortunately, the article did not include the text of their questionnaire, which would have been useful for future researchers. The article did provide a good basic introduction to the subject matter, and described how the library will be moving forward. This is important because it shows the work that the library must do in order to get their brand attached to bioinformatics.

The goal is to support research and the library must make sure it has the tools and can provide the training to support such research.

In “Carving a Niche: Establishing Bioinformatics Collaborations,” Lyons et al. (2006) explored collaboration in depth. The article looked at four institutions and saw how they collaborated with the researchers on their campus. It showed how libraries can be central to the growing collaboration efforts. It emphasized the importance of understanding the structure of the institution in designing the collaboration, because if the players are not known then it is harder to establish any collaborative programs.

“Developing library bioinformatics services in context: the Purdue University Libraries bioinformatics program,” by Rein (2006) does contain a copy of the survey the library conducted during their *Bioinformatics Week 2005*, and so directly informed the design of the current study. Purdue also put a scientist in their library (the author has both a PhD and an MLS). This article reported on two surveys. The first was a survey evaluating the National Center for Biotechnology Information (NCBI) Field Guide, which is a course the libraries offer. The second was a survey on instructional needs at Purdue University, in which the library listed the resources to which it has access and asked the scientists in which resources they desired to have training, and how many people from their lab were likely to attend that training. Rein concludes that, in having a librarian with a PhD in the library, the researchers may decide to approach the librarian more often for help and the PhD may provide her with more ways to collaborate with the researchers than just an MLS could provide.

“A web-based assessment of bioinformatics end-user support services at US universities,” by Messersmith et al. (2006), is an analysis of the library training provided

to scientists. The team of analysts from the National Library of Medicine examined the websites of 239 universities to see if they offered introductory training, advanced training, training only in their libraries, or training in a collaborative fashion. They found that 48% of the universities had bioinformatics centers, but only 33% of them offered training on the tools their scientists needed to use. This was a website analysis, and no users were involved. This methodological point is important to the study because of speculation that more bioinformatics scientists use the library through its web portal than by entering the doors. At the same time the scientists may not be aware of what is at the library. This analysis helped to identify gaps in the web access at different libraries, and in the services a library could be offering.

“A library-based bioinformatics services program,” by Yarfitz and Ketchell (2000), was similar to the Rein (2006) study, but it delved deeper. It pre-dated Dr. Rein’s article by six years and is a foundation piece in the study of bioinformatics in libraries. Yarfitz and Ketchell covered the basic issues that all medical libraries must cover in order to target bioinformatics scientists: a web presence, training courses, a needs analysis, consultation services, skills training, and outreach to the community. The survey, since it was conducted in 2000, does not include all the tools that would be included on a survey conducted in 2007, but it provided a good starting point. The findings showed how the collaboration between the library and the bioinformatics departments and centers on campus can create an open exchange of ideas. Libraries need to make the researchers aware of what services they can provide to them, and these articles volunteer good ways to do that and show what has not been done in previous studies.

The article, “The University of Washington Health Sciences Library BioCommons: an evolving northwest biomedical research information support infrastructure,” by Minie et al. (2006), shows how collaboration between the library and researchers provided resources and training to the researchers in the community around the University of Washington. It has regional and national components and regularly assesses itself as an organization in order to make sure it is doing all it can in order to reach out to the bioinformatics community. It shows a working model of how bioinformatics can work within the framework of a library. In providing examples of self-assessment it provides a clear idea of what the libraries at the University of North Carolina could provide.

The article, “Barriers and enablers to the acceptance of bioinformatics tools: a qualitative study,” examined issues similar to the focus of the current study. Shachak, Shuval, and Fine (2007) conducted qualitative interviews with twelve study participants chosen from a pool of 115 who had previously participated in a study of microarray analysis and primer design workshops. In this study, the scientists were questioned about the workshop and the ease at which they could use the bioinformatics tools presented during the workshop. A qualitative approach was used in order to reach out to the researchers and find out more about their needs, looking closely at the tools the researchers were using and asking probing questions to find out more about their use patterns. This study showed how a library can investigate how a tool is used, and can encourage collaboration by offering courses about different tools and programs.

The third theme found in the literature was the training of librarians in order to support bioinformatics research. The article, “A Model for Training the New

Bioinformatist,” by Lyons et al. (2004), articulates the kinds of training a librarian can undergo in order to aid bioinformatics research. The Eskind Biomedical Library (EBL) at Vanderbilt University designed and implemented a special training course for bioinformatists. Bioinformatist is a term used to refer to librarians who are specially trained in bioinformatics. Lyons et al. evaluated a twelve week course for which fifteen librarians and one library staff person signed up. They were given a pre-test and a post-test and showed significant improvement after the post-test, compared to the pre-test. The results of this study indicated that a librarian could become a specialist and understand enough bioinformatics knowledge to communicate successfully with scientists. The questions involved in the pre- and post-tests are questions that often come up, so it is important for them to also be included in the current study.

The current study was designed to find out whether researchers use the tools, databases and subject expertise the library can provide for them. It hopes to articulate a need for more collaboration, and to find out what bioinformatics researchers need from the library and how the library does and does not satisfy their research needs. It also asks for suggestions for future library services. The University of North Carolina Health Sciences Library wanted to know more about what bioinformatics researchers needed than what could be revealed on a survey. They wanted to have the library be more at the forefront of the researchers’ minds when they began their research and as they worked their way through their projects. In order to achieve these goals, this study was conducted.

Methods

Two articles were especially important in designing the methods for the current study. They were both qualitative information needs studies. In “How do doctors use information in real-time? A qualitative study of internal medicine resident precepting,” Tiburt et al. (2007) observed and recorded residents precepting at outpatient centers affiliated with a Midwestern university. Each attending physician supervised two or three residents and the residents saw one patient every thirty minutes. The researcher observed the clinic setting in half-day increments, recording data on audio tapes and via field notes. The field notes were then transcribed and merged with the audio tape recordings for each half day session. Three investigators read over the transcripts and, through consensus, they developed a codebook with inclusion and exclusion criteria. They analyzed the data and came up with four themes of information exchange behavior: questioning behavior, searching behavior, unsolicited knowledge offering, and answering behavior. This study’s methods are relevant to the current study because Tiburt et al. conducted in-depth interviews requiring qualitative data analysis.

The study entitled, “Information Needs and Information-Seeking Behavior of Primary Care Physicians,” by González-González et al. (2007), analyzed the information-seeking patterns of primary care physicians in Spain and tried to determine their information needs. It was an observational study where the participants were chosen initially through telephone conversations and then were observed and video recorded for four hours. After seeing the patient, each physician was asked to identify all the clinical questions associated with the care of that patient. The questions were categorized in terms of types and topics. At the end of two weeks they called the physicians to find out if there

were any remaining unanswered questions. This study is another example of needs assessment and information-seeking behavior – goals similar to those of the current study. While the methods are not directly transferrable because they were quantitative, the findings showed that the physicians did not use the internet as much as was found in previous studies.

The current study will adapt some of the methods used in these studies, while trying to overcome their limitations. Interviews were conducted with 15 bioinformatics researchers at UNC. Using a qualitative approach it was possible to delve into the bioinformatics researchers' awareness of current library services and what, if anything they wished the library could provide for them. The interview guide allowed the interviewees to open up and articulate their level of satisfaction with the services they have used and their desire to have the library provide new services that will advance their research. The methods used, including the identification and recruitment of participants, the interview structure, and the methods of analysis, are described here.

Identifying and recruiting the study participants

In trying to identify the population of UNC bioinformatics researchers, I examined listings of faculty in many different departments and tried to find the scientists who self-identified as bioinformatics researchers. It was hoped that through the course of this research project, we would find a good population of these scientists who would want to learn more about what the library can do for them, and would be willing to talk about how they already use the library.

The participants were identified and recruited initially through the use of list-servs targeting scientists interested in bioinformatics. They were recruited through the Vector-IT list-serv, the list-serv for the Basic Sciences group on campus, the list-servs for bioinformatics attached to the Health Sciences Library, and list-servs for departments which were considered to be a part of bioinformatics research. The administrative assistants for the departments of Biology, Chemistry, Molecular Biology, Bio-medical Engineering, Epidemiology, Computer Science, Bioinformatics and Computational Biology and the Carolina Center for Genome Sciences were approached and asked to submit the e-mails to their respective faculty and graduate student list-servs. Participants were also recruited via fliers posted around the medical campus at the University of North Carolina at Chapel Hill. There were approximately 150 participants who could have volunteered from the Carolina Center for Genome Sciences alone, without the other institutions or departments at the University of North Carolina that could have provided participants.

The population was further expanded through snowball sampling, asking each study participant or potential participant if they knew of anyone else who might want to contribute to the study. (Babbie, 2004) “Careful purposive selection of informants is almost always the preferred method for qualitative investigation... Informants should be selected based on the information or expertise they can share.” (Friedman & Wyatt, 2006, 278)

Fifteen participants volunteered for the study: three graduate students, four faculty members and two other employees of the university. They came from the schools of medicine and public health, and the microbiology and biology departments. The sampling

was non-random since it is anticipated that few would volunteer for the study, even when asked to volunteer and help their own research practices.

Conducting the interviews

A qualitative interview study is an invaluable way to gather information. The basic interview questions were intended as guides, to inspire dialog with the interviewees, rather than a survey which would elicit dichotomous answers. Depending on the answers given by the respondents the questions may change slightly in order to fulfill the needs of the interview and learn new things to answer the research question. “Qualitative studies are designed so that the topic of interest is considered within a larger context...

Qualitative methods can illuminate the evolution of important phenomena over time...”

(Friedman & Wyatt, 2006, 268) A qualitative study is appropriate because the study is looking at the many ways in which people interact or want to interact with the bioinformatics resources the library supplies.

Each participant was interviewed in his or her office or another mutually convenient location. The interviews lasted for approximately 30 minutes, and covered the ways in which the scientists use the physical library, the ways in which they use the library webpage, the types of contacts they have had with library staff people, and how they understand the future role of the library is in bioinformatics research. The interview guide is attached (see the Appendix). The interviewer audio-recorded the interview, with the participant’s permission, and took notes at the same time.

The interviews were based on main questions, probes, and follow-up questions. “Main questions encourage the conversational partner to expound on your research

concerns, by providing their experiences and perspectives... Main questions are prepared in advance of an interview, though they change as the project progresses.” (Rubin & Rubin, 2005, 200) The second part of the interview is the probing phase. In probing, the interview goes deeper into the details and evidence without interrupting what the interviewee is saying. Probing questions keep the conversation on topic, and are like pieces of a puzzle which will be put together later in the analysis. Follow-up questions are the third part of the interview process. They “pursue concepts and themes that are introduced by the conversational partners and frequently take the interviewer off in new directions.” (Rubin and Rubin, 2005, 200)

These three different forms of questions enable the interview to cover the gamut of any idea that may be relevant in terms of bioinformatics researchers’ needs for library resources and services. It creates a symbiosis between the need of the interviewer to ask questions and the need of the interviewees to volunteer information, since the interviewees may have issues that were not considered when the interview guide was created. It is important to try to cover as many of these issues as possible, to fully answer the research question. Since time is a factor, the probing and follow-up questions were done at the same time as the main interview, and not in subsequent interviews.

Analyzing the interview data

The interview tapes were transcribed and the transcribed interviews were reviewed in order to create a code system by which the interview data could be analyzed. This code system was used to analyze data on the following main themes: Do these

scientists currently use the library? If so, for what purposes do they currently use the library? How could the library be improved to aid in their research?

The data were analyzed by first reading through the transcripts and designing a codebook and then re-reading the data and coding it. The analysis was undertaken in two phases:

In the first, you prepare transcripts; find, refine, and elaborate concepts, themes and events; and then code the interviews to be able to retrieve what the interviewees have said about the identified concepts, themes, and events. In the second phase several paths are followed. You can compare concepts and themes across the interviews or combine separate events to formulate a description of the setting. In doing so, you seek to answer your research question in ways that allow you to draw broad theoretical conclusions. (Rubin and Rubin, 2005, 201)

After the transcripts were created, they were read carefully and any ideas of interest were noted. They were then compared to the next interview, and ideas of interest were noted. After a few transcripts the ideas of interest began to coalesce into categories. Those categories became the codebook for analysis. Qualitative data analysis is not about mere counting, or providing numeric summaries. Instead the objective is to discover variation, portray shades of meaning and examine complexity.

In an interview, an issue can be explored through the give and take of questions. The interviewee can express information more readily than in a survey because they are volunteering it, with some sort of explanation or rationale behind it. The explanation becomes as important as the answer, since it brings out the complex ideas behind simple answers to questions.

Results and Discussion

The results of the interviews will be presented and discussed here, in the order in which they were structured in the interview guide. Where appropriate, a quantitative summary of the responses is provided; in most cases, the results are presented qualitatively.

Characteristics of the study participants

In order to get a clear idea of the distribution of the participants they were asked how many years they had been working or studying at the University of North Carolina. Most had been at UNC for four to seven years. The complete results are presented in Table 1.

Table 1. Experience at UNC

Number of Years Working or Study at the University of North Carolina	Number of People
1 year	1
2.5 years	1
3 years	1
3.5 years	1
4 years	3
5 years	2
6 years	2
7 Years	2
8 Years	1
13 Years	1
Total:	15

Of the fifteen study participants, 27% claimed that Bioinformatics is their primary discipline and 73% identified something else as their primary discipline (see Figure 1). This is important because it helped the respondents to frame their ideas of whether or not they were bioinformatics researchers. Many of them considered themselves first

biologists, physicians or epidemiologists who worked with bioinformatics tools, and not bioinformatics researchers. As the fields become more interdisciplinary, the number of people who will consider themselves bioinformatics researchers will hopefully increase, as traditional disciplines disappear and non-traditional disciplines with cross-disciplinary bents emerge.

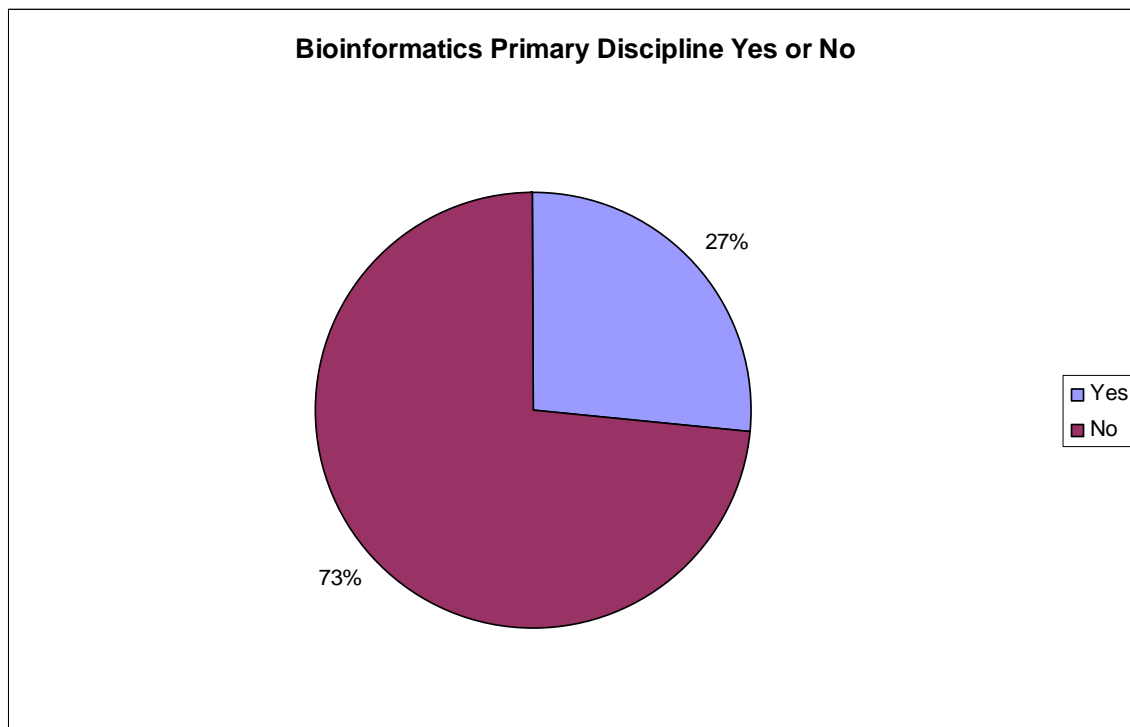


Figure 1. Primary discipline

Career status is important because it shows whether or not participant was a graduate student, researcher, faculty member or post-doctoral student. It was expected that the graduate students would have less interaction with bioinformatics resources since they have been in the field less than the researchers, post-doctoral students or faculty members. The distribution of the study participants across the different status levels is shown in Figure 2.

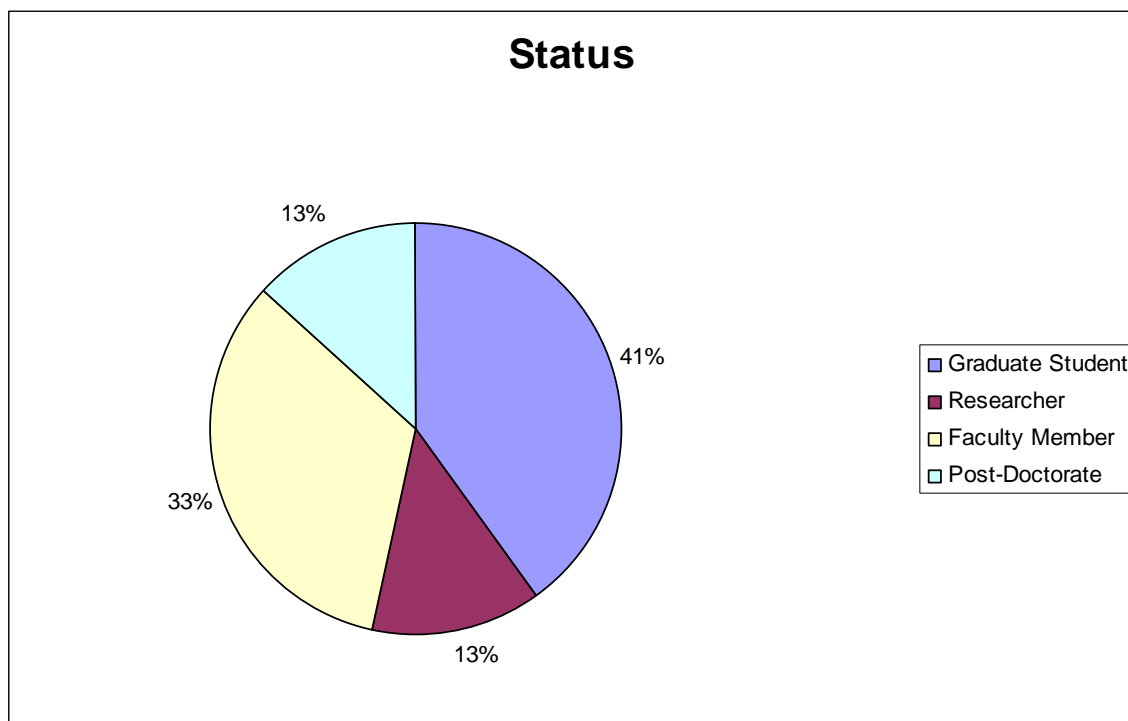


Figure 2. Status of participants

The majority of the participants were from the department of Microbiology. This is because an entire laboratory volunteered for the study (see a discussion of this point in the later section on limitations of the study). Medicine was also a large group because several of the interviewees were medical doctors or worked in the hospital. These large groups from one or two specific disciplines could lead to certain biases in this study's findings. A summary of all the departments represented is shown in Figure 3.

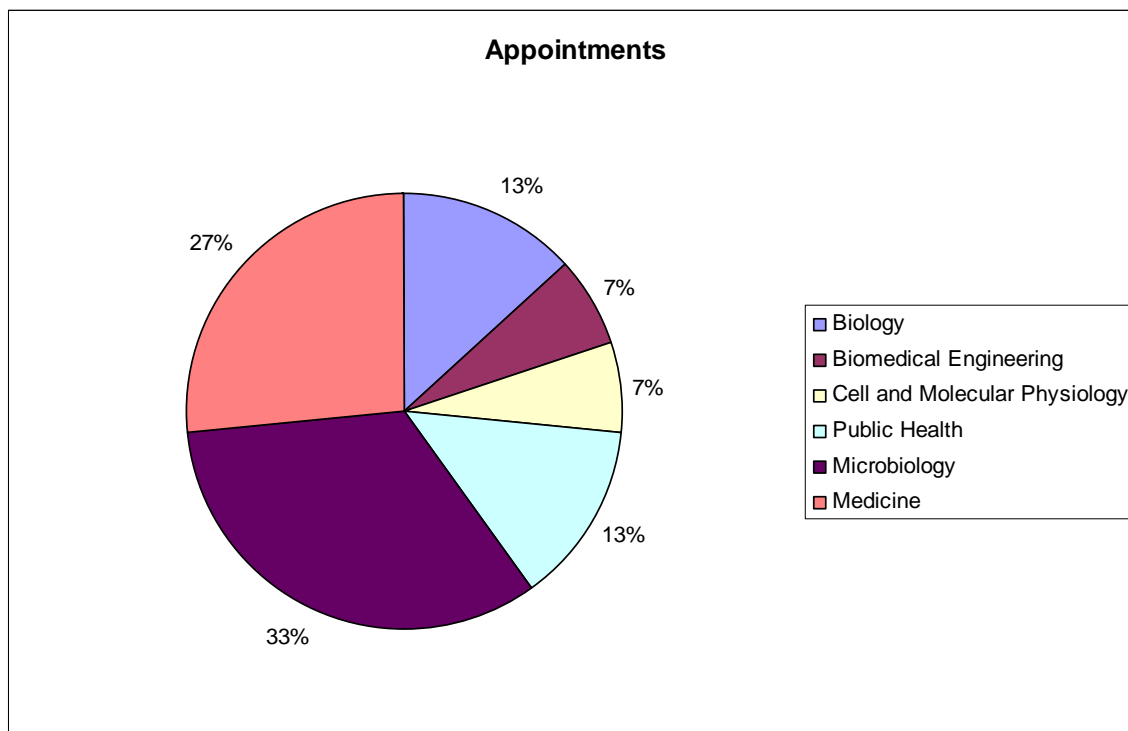


Figure 3. Departmental appointments of study participants

The physical library

The majority of the participants had visited the Health Sciences Library most recently (see Figure 4). This makes sense since the Health Sciences Library is the UNC library which is located closest to their labs and offices. They rated their experience on their most recent visit to the library very positively, with eight participants rating it with a five and six participants rating it with a four, on a five-point scale.

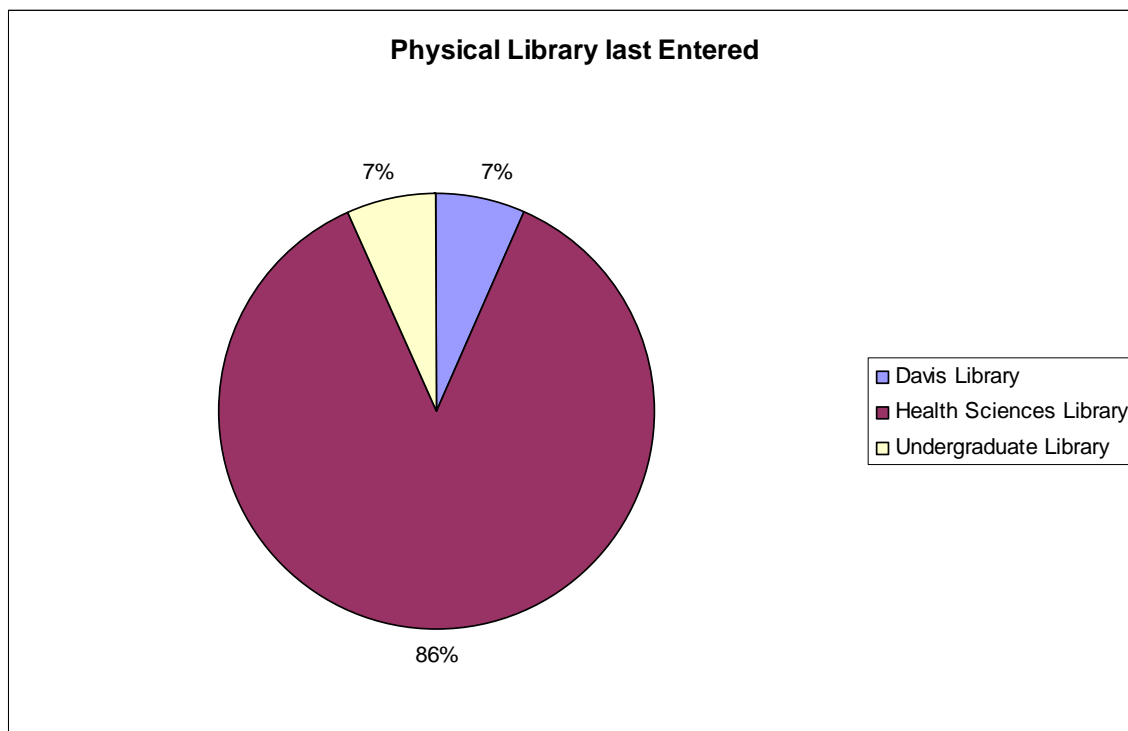


Figure 4. Physical library last visited

The future of the physical library. The future of the physical library, according to these bioinformatics researchers, depends on where they are in their programs. The graduate students, on average, had entered the library more recently than either post-docs or researchers. As with every other study, they had varying reasons for entering the library, ranging from needing to get a book or pay a fine to needing a quiet place to study.

They also had varying ideas on how the structure of the library could be adapted to better serve their needs. According to one participant, "I think I would change the number of group study rooms or the group study policy because I would like an increase in them. I think that sometimes, for students who are easily distracted, [group study rooms] are great to use on an individual basis to keep out some of the noise. I like the

quiet floor but there is not always space available so sometimes you go to a group study room, and you feel guilty using that group study room, you know, if it is just you.”

Several participants felt that the library was important when it came to being a place to study. They spoke about it being a quiet space. “In the past, whenever I was in school, it was nice to have a place to work in that was quiet, and that was always the library at the schools I was at. So I think it is important for studying.”

A lot of the participants spoke of using the library as a gathering place. They enjoyed the idea of the information commons where, like the village commons of old, there was a place for people to meet and, just as medieval villagers traded gifts, researchers could trade information. “I think a gathering place, sort of a gathering place where intellectual discussion and research can occur, and things like this class that I taught.”

As scientists, the participants often did not have any need for a physical library. They felt that everything they needed or would need was already online. “Aside from what I just said, as a place where people can go and it is a dedicated work area, I think that there is not much role for the physical housing of information in pages, just from my perspective.” Still, one participant articulated the need for a physical library, in spite of the push for journals to all be online:

That you can go to actually see books and journals and seeing first hand what is present. We are so increasingly converting to electronic information, but you need that hard copy back-up. So my prayer is that we don't convert totally to electronic because that makes us more vulnerable to some major mishap with electronic files. It also has to be increasingly that we'll save ourselves and the university a fortune if a reasonable percentage of the journals are available primarily in electronic form. What I'm projecting is that the university would keep a hard copy or a couple of hard copies of something, but most of what would be distributed among people would be the electronic forms and the hard copies would be back-ups.

When asked about the physical library, another idea that was brought up was textbooks. Textbooks, in either physical or electronic form, were important to many that were interviewed. This idea was put forth that the physical library should “provide to the campus community up to date resources. And if we’re talking about the physical building, just being a source of the current textbooks that all the courses on campus may require for their coursework, so that students can have access to them whether or not they can afford their own texts and stuff like that.”

Journals and a journal repository was another idea they thought was important. Though they wanted all the journal articles online, they understood that may never happen. “I guess a repository for academic journals that may or may not be available on the internet and basically I guess just provide a training service for just the basics for one how to use the library effectively and also you know generally source the information that they need whether it is from the physical library or you know not.”

One participant had a different take on the physical library. He believed that there should be “not just one library, but several locations so that it is convenient for many people on campus. So maybe smaller sites but all, of course, with internet access (wireless) so that we could do research – our searches online [and] downloading information. So not really that different than how it is now, except that it could be more accessible to more people.”

Two participants differed with that assessment, believing that there should be one central library so they would not have to go all over campus to get the resources they needed for their research. One stated that “it would be nice to have everything in one central library location, because I don’t appreciate having to go to -- let’s see I’ve been to

the zoology, botany, math/physics, health sciences and undergraduate [libraries] to get books that I need.” Another claimed: “I am still concerned about the libraries in UNC. [It] seems they are kind of separated, individually, so sometimes it is just hard to find a particular resource in a single place. You just, maybe, need to go to several places to get what you need.”

In summary, the majority opinion was that there may not be a need for a physical library if everything is online. In spite of this widely-held view, individual participants did mention the need for the library to store journals as a back-up for the electronic copies and the need for the library to provide access to course textbooks. In addition, the library was seen as a place for quiet study, as well as group work and interactions with research colleagues.

The future of the physical library in bioinformatics. Overall the participants did not think there was much of a role for the physical library in aiding bioinformatics research. The research materials they need are mostly online, so the role of the physical library had not been something that they considered. “I never thought they would have to support this research. I never thought they would have to be involved with it” was a common response from the researchers. The participants did not like that the older journals were not online, and that they were required to go into the library to find a key piece of literature. They appreciated being able to access articles from their personal computers and read them at their leisure rather than having to make time to visit the physical library.

Among other responses to the physical library aiding in bioinformatics research included the idea of the library as a spot to hold training courses and as a place to house

information. One researcher thought that “the physical library can in some sense house the information, whether it’s on a server or whether it is in literature, whatever people prefer.”

Another participant thought that the physical library was important for having meetings between different members of the scientific community, and that “the role [of the library] is to provide ways to quickly bring lots of different sources of information together, that all are related by some common mechanism, even if it isn’t just about text.”

Many participants were unaware of what the physical library could provide for them in terms of supporting their bioinformatics research. As one noted, “I don’t have a clear idea about what I would need from the library about bioinformatics.” These library users need to be educated as to what the physical space could do for them, through examples of what other libraries have done to support the bioinformatics research on their respective campuses.

The role of the virtual library

The next section of the interview dealt with the participant’s most recent use of the library website or web-based services provided by the library, i.e., their use of the “virtual” Health Sciences Library. They usually accessed the library’s web pages through the Health Sciences Library page, or a derivative page. During the interview, participants were asked to rate the quality of their recent experience with the library website. Most participants were quite positive in their ratings (three rated it as four, and eleven rated it as five on a five-point scale), but one person was less positive, rating it as two.

The use of the virtual library for promoting bioinformatics research seemed tied to the regular use of the virtual library by the participants. The participants did not see much

of a difference between the sources the library could provide and the sources the library already provided to them. One researcher wanted the most basic of library services applied to the future of the virtual bioinformatics library, saying, “I guess I just wish there was better indexing of the services that are available. It is difficult to find out what is available sometimes and with the search engine there.”

When asked to expand on their ideas of what they could have in a virtual library, the participants provided some interesting suggestions. A few of them wanted tutorials in order to learn how to use the tools more efficiently. One felt that there was a lot of analysis software in existence, but was not sure how to use it with his research. He claimed that, “I think it would be good if the library staff had some tutorials on how to use some of the software that is available for these kinds of analyses. I don’t know if that is a statistician’s job, though, because some of them could be really complex pieces of software. There is a lot of free software out there that I use but I don’t even know... like, some of them, I don’t know how to use it because there’s no good manual. And it would be nice if somebody could give you a tutorial.”

Another researcher also praised the idea of tutorials. He wanted them online, and restricted to the UNC campus, since he does not have the time to sit through a class:

Tutorials would be great. Sort of like information sources. They might already exist, and I’m just not using them. Maybe like the training classes they had. It would be more useful if they were access restricted to UNC students only; but tutorials, online, where you don’t have to go and sit two hours in a class. That’s probably why people would avoid it, that’s why I avoid it. I don’t have two hours to sit in a class when I might only use ten minutes of applicable information.

Researchers want to have easier searching tools, and want the virtual library to be more streamlined so they do not have to jump between databases. They want online tutorials and a more comprehensive way to identify the online databases they need for

their research. This is an important place where the library can advance and support research because, as one researcher stated, “In terms of accessing this exponential growth in number of genomic data bases...It’s just skyrocketing. The library will have a valuable role in helping researchers access this information that’s out there.”

The role of the librarian

The last portion of the interview focused on the participant’s most recent interaction with a library staff member. The participants preferred to find information themselves rather than to ask for help from a library staff person, but when they did ask for help, they told poignant stories of what they learned from the librarians. They did not identify who were professional librarians and who were other staff members. Participants’ ratings of the quality of their recent interactions were mixed: nine rated it as a five, two as a four, and two as a three (with two not providing a rating). These varied ratings were due in part to what the participant needed and in part to their impressions of the staff person’s level of service.

The participants had differing ideas about the future role of the librarian. A lot of them considered the librarian’s role to be fine the way it is. Others wanted it to change. One graduate student predicted the librarian’s role would become one of a data manager, with no physical library existing anymore. Their demeanor, when asked about the future role of librarians, generally indicated that they were surprised that they would be questioned about it. They seemed to understand the role of the library, but not the role of the librarian. Unfortunately a few of the interviewed responded with, “I don’t know.” But others had some interesting responses. As with the responses about the last time they

went into a library, most of the comments were specifically about the Health Sciences Library at the University of North Carolina and its staff.

The participants wanted there to be more advertising of the services that librarians could provide, in relation to their research. These services and facilities may not be familiar to all the researchers:

I do wish they advertised more about what was available because you know a lot of students don't even know about things like the media kitchen or... all the other services that are offered. And I think people just assume they know how to use search engines, but they really don't, and they also don't know all the search engines that are out there.

Training was another role that the librarians could take on that was mentioned by the study participants. They felt that, in five to ten years,

more and more librarians will probably take on more of a training role, in that they'll need to be familiar and keep abreast in advances and database searching and whatever tools are available, whether they are in informatics or not, that graduate students may need for their work. They may not necessarily have to have in depth knowledge of any particular tool but just being generally aware of what's out there and available so that they can guide people of where to go next or where to look.

The participants also believed that they would learn more if they had a librarian who specialized in their subject areas. "I guess if I had someone that was, like, within my area of research, that could point me in the right direction of information or articles I am interested in, sort of like my own virtual librarian. That would be a nice thing." Along similar lines, the participants believed that librarians are "gonna help investigators and people to connect with what is available in the literature, and whether the literature is electronic or in hard copy. So I think the librarians in the future are gonna probably be increasingly versatile in computers and how to access databases."

The future role of the librarian in bioinformatics research. Librarians need to publicize their services more, in order to show researchers what skills they can bring to the research arena. Some of the participants were aware of the skills librarians had, and what they could bring to the future of bioinformatics, but the majority of them could not identify why a librarian would be involved in bioinformatics. “I don’t know. I haven’t used a librarian. Probably to sort of facilitate these sorts of things; I’m more interested in the online part of it. Probably they [could] facilitate those sorts of tools.”

In thinking about the role of the librarian in bioinformatics, researchers considered the librarian’s role to be similar to what they do in other fields. As one participant said,

It is the same as all the other fields. It is enabling people to cross fertilize across fields, to get access to the original documents whether they are in electronic form or hard copy, and learning how to use these wonderful databases, but people have to be trained in them so they can really use them proficiently. I think the librarians’ role will increasingly be to train people to access the electronic forms of the documents and the electronic databases.

Several of those who were interviewed were just beginning to delve into the world of bioinformatics. They were not certain where they could receive more information and training about the discipline, and believed that the librarian might be a resource for obtaining more information. They mirrored the words of one researcher who stated that, “I guess [librarians are] a resource to know what programs exist, or helping [us learn about] these programs. Sometimes I’m not sure where I’m supposed to go to. I’m not sure how to learn new things or how to get access to software. Unless I come across somebody who already knows it, I wouldn’t know where to go ... so that might be a good resource.”

Differences in the views of different subgroups of participants

There were a few discernable differences between the different subgroups included in the study sample. The staff people and faculty were less likely than the graduate students to have recently used the physical library, but all groups used the virtual library extensively. The graduate students and staff members had more insight into the way they thought the library should change in the next five to ten years. They wanted there to be more quiet study spaces and group study rooms in the physical library, and they wanted the online journal collection to encompass all of the articles they would need to satisfy their research needs. The graduate students and staff members were less certain about what a librarian did, so being asked about the changing role of the librarian was one section of the interview to which they were not able to respond as well as the other sections.

The faculty members had more to say about the changing role of the library in the field of bioinformatics. They saw the library as a place where there could be more training and a place which could point them towards more resources. One subject volunteered that he saw the role of the bioinformatics librarian as one of data manager, and wanted that person to design a program that would mine journal data so he could access exactly what he wanted, and would not have to dig through journals in order to find the pertinent articles.

Limitations of the study

As with any empirical work, this study has some limitations. The study was conducted over a short period of time (from January until April 2008). In having the

study compressed into such a small amount of time, the number of participants willing to volunteer became an issue. According to the websites for the Carolina Center for Genome Science, there are fifty faculty members belonging to the center, all of whom potentially could have volunteered for the study. Most of the faculty members have websites which list their lab members. The listed labs contain approximately 150 people. Having 15 people in the study, means that it is approximately one tenth the size of the potential pool from the Carolina Center for Genome Sciences. This implies that the results from the study may not be generalizable to the entire population of bioinformatics researchers at UNC. The sample was likely to be biased towards those who use the library, those who like libraries, or those who are already aware of the benefits the library could provide for them. It did not specifically target those who do not use the library or those who do not know what the library can do for their research, so those groups are likely to be under-represented in this sample

Conclusion

The results from this study provide some initial indications of what this specific population – bioinformatics researchers – needs from the academic library. Prior work has not delved this deeply into what resources and services are needed by this population, and it is important to acquire this level of understanding in order to adjust the library's offerings in order to meet those needs.

As libraries become more targeted toward specific populations, the targeting of the collection becomes increasingly important and studies like this need to be maintained to confirm the collection of bioinformatics materials and databases is sufficient to aid this

population. The participants in this study indicated that libraries should forge ahead in their efforts to make research materials available online, but should also be cautious in maintaining secure backups of important materials in hard copy. While holding of the collection is becoming a less important role for the library, the physical library as a place to study and to interact with colleagues, continues to be a critical role for its users. In addition, an important role that can be played by librarians is that of trainer and guide to the use of new research tools and databases.

The reference librarian can no longer sit behind the desk waiting for the patrons to come in and talk to him or her. Instead, librarians must go out into the population and find patrons. In seeking these patrons, who either currently use the library or, hopefully, will start to use the library resources, this study provides data through which the library can target bioinformatics scientists and satisfy their information needs.

Currently the UNC bioinformatics community is widely dispersed, and there is little known about which members of the community know what the library can supply, or in which research areas their colleagues in other departments at the University of North Carolina are engaged. The hope is that the library can become a connection between departments and individual researchers, thereby creating more collaboration between scientists in bioinformatics research.

Bioinformatics is a new field which needs more exploration. There is no consensus, yet, as to what disciplines it encompasses. The library needs to be at the forefront of this discipline since it can be a place where the different disciplines can freely exchange ideas. This study will start a dialog at UNC, between the library staff and the scientists and graduate students conducting bioinformatics research. It will hopefully

create a greater symbiosis between the two, as they use each other to advance the field of bioinformatics. The future of bioinformatics is a future of collaboration in which the library hopes to partake. This study has endeavored to be a stepping stone toward a collaborative community where scientists and the library work together to further scientific research in the cross-disciplinary and interdisciplinary field of bioinformatics.

Bibliography

- Babbie, E. *The Practice of Social Research*. 10th ed. Belmont CA.: Wadsworth-Thompson Learning, 2004.
- Chattopadhyay A, Tannery NH, Silverman DA, Bergen P, Epstein BA. Design and implementation of a library-based information service in molecular biology and genetics at the University of Pittsburgh. *Journal of the Medical Library Association*. 2006; 94(3): 307-13.
- Claverie J, Notredame C. *Bioinformatics for dummies*. New York: Wiley Pub., 2003.
- Friedman CP, Wyatt JC. *Evaluation Methods in Biomedical Informatics*. 2nd ed. New York: Springer Science and Business Media, 2006.
- Geer RC, Rein DC. Building the role of medical libraries in bioinformatics. *Journal of the Medical Library Association*. 2006; 94(3): 284-5.
- Geer, RC. Broad issues to consider for library involvement in bioinformatics. *Journal of the Medical Library Association*. 2006; 94(3): 286-98.
- González-González AI, Dawes M, Sánchez-Mateos J, Riesgo-Fuertes R, Escortell-Mayor E, Sanz-Cuesta T, Hernández-Fernández T. Information needs and information-seeking behavior of primary care physicians. *Annals of Family Medicine*. 2007; 5(4): 345-52.
- Giuse NB, Huber JT, Giuse DA, Kafantaris SR, Stead WW. Integrating health sciences librarians into biomedicine. *Bulletin of the Medical Library Association*. 1996; 84(4): 534-40.
- Helms AJ, Bradford KD, Warren NJ, Schwartz DG. Bioinformatics opportunities for health sciences librarians and information professionals. *Journal of the Medical Library Association*. 2004; 92: 489-93.
- Lynch C. Medical libraries, bioinformatics, and networked information: a coming convergence? *Bulletin of the Medical Library Association*. 1999; 87(4): 408-14.

- Lyon J, Giuse NB, Williams A, Koonce T, Walden R. A model for training the new bioinformaticist. *Journal of the Medical Library Association*. 2004; 92(2): 188-95.
- Lyon JA, Tennant MR, Messner KR, Osterbur DL. Carving a niche: establishing bioinformatics collaborations. *Journal of the Medical Library Association*. 2006; 94(3): 330-5.
- MacMullen WJ, Vaughan KTL, Moore ME. Planning bioinformatics education and information services in an academic health sciences library. *College & Research Libraries*. 2004; 65(4): 320-34.
- Messersmith DJ, Benson DA, Geer RC. A web-based assessment of bioinformatics end-user support services at US universities. *Journal of the Medical Library Association*. 2006; 94(3): 299-305.
- Minie M, Bowers S, Tarczy-Hornoch P, Roberts E, James RA, Rambo N, Fuller S. The University of Washington Health Sciences Library BioCommons: an evolving northwest biomedical research information support infrastructure. *Journal of the Medical Library Association*. 2006; 94(3): 321-9.
- Norman F. Genetic information resources: a new field for medical librarians. *Health Libraries Review*. 1999; 16: 15-28.
- Rein DC. Developing library bioinformatics services in context: the Purdue University Libraries bioinformaticist program. *Journal of the Medical Library Association*. 2006; 94(3): 314-20.
- Rubin HJ, Rubin IS. *Qualitative Interviewing: The Art of Hearing Data*. 2nd ed. Thousand Oaks CA: Sage Publications Inc., 2005.
- Shachak A, Shuval K, Fine S. Barriers and enablers to the acceptance of bioinformatics tools: a qualitative study. *Journal of the Medical Library Association*. 2007; 95(4): 454-8.
- Tilburt JC, Goold SD, Siddiqui N, Mangrulkar RS. How do doctors use information in real-time? A qualitative study of internal medicine resident precepting. *Journal of Evaluation in Clinical Practice*. 2007; 13: 772-780.
- Yarfitz S, Ketchell DS. A library-based bioinformatics services program. *Bulletin of the Medical Library Association*. 2000; 88(1): 36-48.

Appendix:

Interview Guide:

Preliminary demographic questions:

- In which UNC department do you work/study?
 - What is your status there? (Faculty, researcher, post-doc, graduate student)
 - Would you consider your primary discipline to be bioinformatics? If not, what is your primary discipline?
 - How many years have you been working or studying at UNC?
1. Think about the last time you entered a library on [the UNC-CH] campus.
 - Which library was it?
 - When was this visit?
 - Why did you visit the library?
 - Did you get what you needed on this visit?
 - Rate your experience on a scale of one to five, one being the most negative and five being the most positive. Why did you choose this value?
 2. Think about the last time you accessed the library's website.
 - Which page did you visit? (give options: HSL home, Davis home, Clinical Ref, Pharm Ref, Other)
 - When was this visit?
 - Why did you visit this page?
 - Did you get what you needed on this visit?
 - Rate your experience on a scale of one to five, one being the most negative and five being the most positive. Why did you choose this value?
 3. Think about the last time you contacted a library staff person.
 - Who did you contact? (give options: Ask A Librarian, specific contact (name?))
 - When was this?
 - How did you contact this person? (options: email, phone, chat, in person)
 - Do you know if this person was a librarian?
 - Why did you contact this person? (some options: research question, access, circulation, reserves, room issues, copying/printing, ILL, etc.)
 - Did you get what you needed on this visit?

- Rate your experience on a scale of one to five, one being the most negative and five being the most positive. Why did you choose this value?
4. In your opinion, what is the ideal role of the library on an academic research campus such as UNC-Chapel Hill's in 5-10 years?
 - Physical
 - Virtual
 - Librarian
 5. In the field of bioinformatics (genomics, genetics, proteomics, molecular biology, computational biology) what do you see as the ideal role of the library in supporting these fields of research?
 - Physical
 - Virtual
 - Librarian

Thank you for your time. I am interviewing as many bioinformatics researchers as possible. I have sent an email message to several bioinformatics-related listservs to recruit study participants. Would you provide me with the names of any bioinformatics researchers who should be included in this study that I might have missed?