University of Iowa Libraries Staff Publications

9-1-2017

Public Health Research Practices at the University of Iowa

Janna C Lawrence
University of Iowa

Chris Childs
University of Iowa

Public Health Research Practices at the University of Iowa

Janna C. Lawrence, Deputy Director
Chris Childs, Education and Outreach Librarian
Hardin Library for the Health Sciences, University of Iowa, Iowa City, IA

September 2017

Introduction

The University of Iowa's College of Public Health was founded less than 20 years ago in 1999, but public health at the University of Iowa reaches back another century. The State Hygienic Lab has been located at the University of Iowa since 1904, and the forerunners to the College's Department of Occupational and Environmental Health and the Health Administration graduate program were formerly based in the Carver College of Medicine. The College of Public Health currently has approximately 80 full- and part-time faculty in five academic departments: Biostatistics, Community and Behavioral Health, Epidemiology, Health Management and Policy, and Occupational and Environmental Health.

The University of Iowa Libraries was approached by Ithaka S+R, a not-for-profit research and consulting service, about taking part in a study is to examine the research practices of the academics at the University of Iowa's College of Public Health in order to understand the information resources and services these faculty members need to be successful in their scholarship. The study at the University of Iowa was part of a broader suite of parallel studies by librarians at colleges of Public Health at seven U.S. universities. The same set of questions and methods for performing the research and analyzing the resultant data were used at each university, and the study was vetted by each university's Institutional Review Board. Ithaka S+R provided guidance on research methodology and data analysis for the project and produced a comprehensive report analyzing the anonymized data from all seven universities.

The study of Public Health is one of a series of studies from Ithaka S+R on the research support needs of various disciplines. Previous studies have examined needs in the fields of agriculture, art history, chemistry, history, and religious studies. Ithaka S+R has chosen to explore fields which have not previously been studied in depth.

This paper explores the findings of the interviews performed at the University of Iowa by librarians from the Hardin Library for Health Sciences. It will focus primarily on data and publishing, two broad concepts that were among the most frequently mentioned, and on research focus, which helps define the field of Public Health. It will also examine challenges to research expressed by the faculty members interviewed.

Methodology

Chris Childs, Education and Outreach Librarian and liaison librarian to the College of Public Health, and Janna Lawrence, Deputy Director of Hardin Library, interviewed eight faculty members from the College of Public Health. Seven of the faculty members held the rank of Professor, while one was an Assistant Professor. Faculty from all five academic departments in the College were interviewed, with two

interviews each from Biostatistics, Community and Behavioral Health, and Epidemiology, one from Health Management and Policy, and one from Occupational and Environmental Health. Interviews were held between January 20 and March 8, 2017, and ranged in length from 18 minutes, 47 seconds to 1 hour, 12 minutes, 57 seconds, with an average of about 41 and a half minutes.

Interviews were then coded by the interviewer, using an open coding technique without a controlled vocabulary. After coding, the interviewers met to standardize the codes used. For example, one interviewer used the code **Dissemination Process**, while the other used **Dissemination Practices**; in this case, all instances of **Dissemination Process** were changed to **Dissemination Practices**. The resulting standardized list included 303 codes. The most used code was **Challenges**, with 36 occurrences, followed closely by **Data Collection**, with 35 occurrences. Table 1 below shows the 13 codes which each had 10 or more occurrences. Sixty percent of the 303 codes (184) were used only once, and another 20% (61 codes) were used just two times.

Table 1: Codes with 10 or more occurrences

Code	Occurrences	Code	Occurrences
Challenges	36	Open Access	12
Data Collection	35	Publishing Practices	12
Collaborations	27	Publishing	11
Data Management	17	Data Analysis	10
Data Storage	16	Data Elicited from Research	10
Keeping Up with Trends	13	Dissemination Practices	10
Journals	12		

After standardizing the codes, the interviewers then grouped together similar codes under broad concepts such as *Data*, *Challenges,* and *Research Focus*. Note that throughout the paper, words used as codes will be in **bold** type, while words used as broad concepts are in ***bold, italicized*** type.

While looking at the codes in this manner, the interviewers discovered a few codes which seemed very similar and perhaps referred to the same concept, but which had not been standardized. For example, Interviewer 1 had used the code **Data Elicited from Research** ten times, while Interviewer 2 had used the term **Data Collection** for the same concept. However, the first interviewer had also used the term **Data Collection**, and felt that there was a subtle difference between the two concepts, with **Data Elicited from Research** referring to data collected directly from the researcher's own research subjects, while **Data Collection** referred to data obtained from other sources. Therefore, both codes – **Data Elicited from Research** and **Data Collection** – were retained, although Interviewer 2 did not attempt to reconcile the difference in their interviews by adding the code **Data Elicited from Research**.

Results

Two hundred fifty eight of the 303 codes (85%) were grouped into 14 broad concepts. Of the 45 codes that were not grouped, 33 of them had only one occurrence and did not readily relate to other codes, while codes like **Communication** with 6 occurrences and **Follow Up** (4 occurrences) seemed to stand on

their own. Table 2 shows the broad concept terms, the number of codes grouped in each, and the total occurrences of those codes in the interviews.

Table 2: Broad concepts

Broad Concept	Number of Codes in Concept	Total Occurrences of Codes
Academia	13	19
Challenges	18	75
Collaborations	26	60
Community	3	8
Data	44	163
Funding	3	11
Information	17	26
Public Health (field)	3	13
Publishing	32	104
Research Focus	31	39
Research Methods	35	66
Social Media	2	5
Study Participants	5	7
Technology	25	41

The interviewers have chosen to focus on a few of the concepts, or grouping terms, listed above in this paper. Because the broad concept of *Data* included the most number of codes (44) as well as the most total occurrences of those included codes (44), it was the first concept chosen. *Publishing* was the second concept selected; it was second in total occurrences of codes (104) and third in the number of codes included (32), and was also a topic of great interest to the librarian-interviewers. The third concept in number of occurrences (75) was *Challenges*, and was a theme that was woven throughout the interviews, even though the challenges discussed were not always coded as such. Finally, the interviewers wanted to look more closely at the codes groups under the concept of *Research Focus*, since that topic seemed integral in understanding Public Health research. Because of this, it was the first concept examined.

Research Focus

The field of Public Health is incredibly wide-ranging, with researchers who perform clinically-related research, others who do research that is more social science-related, and others, primarily biostatisticians, who may "hire out" to groups who need their skills. Additionally, individual researchers may work in a broad area, but have numerous diverse projects. Most research in Public Health is interdisciplinary, with many faculty working with researchers outside of Public Health.

Research Focus is not one of the larger grouping terms, encompassing 31 codes but only 39 occurrences, but understanding the extensive and varied research foci of Public Health professionals is instrumental in understanding Public Health. As Table 3 shows, most of the codes used delineate areas of research.

Table 3: Codes grouped under the broad concept *Research Focus*

Code	Occurrences	Code	Occurrences
Research Focus	7	Literacy	1
Behavior	2	Mapping Work	1
Mathematics	2	Medicare	1
Bullying	1	Medicare Part D	1
Child Abuse	1	Motor Vehicle Crash	1
Cyberbullying	1	Older Adults	1
Disability	1	Pedestrians	1
Elder Abuse	1	Pediatrics	1
Environmental Health	1	Pesticide Users	1
Firearm Injuries	1	Physical Activity	1
Gerontology	1	Preventive Health Care	1
Health Care Delivery System	1	Prevention	1
Health Care Services	1	Prevention Research	1
Health Insurance Marketplace	1	Rural Areas	1
HPV Related Cancer	1	Schools	1
Injury and Injury Prevention	1		

A repeated focus for the researchers in this cohort, although not obvious from the code assignments, was rural health. One researcher directs projects in the Prevention Research Center for Rural Health. This includes a project partnered with a group in Ottumwa, Iowa, focused on increasing physical activity. She also heads community outreach activities for the Environmental Health Sciences Research Center and the community engagement core for the university's Institute for Clinical and Translational Science.

Another researcher declared that "the general focus of everything that I do is related to health care services in rural places." Projects include work on sustainability of rural pharmaceutical services. This researcher works with a team at another Midwestern university on project pertaining to the insurance marketplace and health reform over the years. He works with the Rural Resource Center at the local Veterans Administration hospital and on a project related to Medicaid in Iowa.

One researcher is the Director and Principal Investigator for the statewide cancer registry, which collects information on every Iowan diagnosed with cancer over the last 40 years. Many of the projects coming out of that resource are related to the rural nature of Iowa, including one that studies the relationship between types of farm pesticides and cancer and another that is looking at the workforce of gynecologic oncologists in the upper Midwest. However, other current projects study the impact of HPV vaccination and improvement of cancer registries

Another repeated theme was driving studies. Since the University of Iowa is home to the National Advanced Driving Simulator, this focus is very logical. One researcher who does primarily interventional studies and works with researchers from Engineering, Education, Computer Science, and the Arts, in

addition to Public Health and the College of Medicine, looks at "individual, family and community characteristics as to who might be at high risk for a motor vehicle crash" as well as studying groups such as "novice drivers, bicycle rider exposure, and farm equipment on the road." Another researcher mentioned that he has worked on driving studies with the Department of Neurology from the College of Medicine, studying driving in neurologically impaired individuals.

Social relationships and their implications on people's health are studied by a researcher who is part of the campus-wide gerontology initiative, which involves faculty members from the Colleges of Nursing, Business, and Engineering, as well as from the Liberal Arts departments of Anthropology, Sociology, and Computer Science.

The common theme of "prevention and disability" is one researcher's work extends to clinical trial analysis working with College of Medicine faculty in the Departments of Orthopaedics, Neurosurgery, Surgery, and Emergency Medicine. This researcher also works with state-wide "disease registries in stroke and trauma for statewide surveillance and care."

Two faculty members from Biostatistics were interviewed. Both spoke of doing two types of research, one clinical and one methodological. In clinical research, biostatisticians might formulate the design for collecting data, then analyze the collected data. However, they both also do research in statistical methodology involving algorithms and mathematical modeling. One explained that, as biostatisticians, "we do have specific requirements as we go through the tenure ranks, of doing a certain amount of that [statistical research]. We're allowed to kind of specialize and have either a concentration in methods or in collaboration as we get further in our careers. But especially as we're going from assistant to associate professor, as we're going for tenure, we want someone to have a good mix of both of those kinds of things." He also mentioned that he believes he has "collaborated with over 20 different medical departments here on the health sciences campus, with publications with virtually all those." Nevertheless, he feels that he has developed specialties within these collaborations, working on the previously-mentioned driving studies with the Department of Neurology, elder abuse with the Department of Family Medicine, and cardiovascular research with faculty from Family Medicine, Pharmacy, Pediatrics, and Neurology.

The other stated that biostatisticians "are expected to be involved in both types of research... In the Department the focus is mainly on disciplinary methodological research…. [W]ithin the domain of interdisciplinary collaborative research, I tend to have three areas of focus. I have an appointment in what's called the Injury Prevention Research Center [which] covers the broad range of injuries and injury prevention. Some off the studies that I've currently been involved in deal with trying to prevent high school and middle school bullying and recently we just published a paper that looks nationwide at bullying legislation and tries to determine how effective that has been in reducing the incidence of bullying. I also work with a collaborative team on the epidemiology of infectious diseases, tracking infectious diseases which ones are on the rise. How different disease incidence patterns might be related to one another and then finally I work in the Colleague of Dentistry on this large scale, longitudinal study for the state of Iowa that's been ongoing since the early nineties called the Iowa Fluoride study."

In addition to their research efforts, several faculty mentioned teaching and administrative activities. Although these activities might seem to distract from research activities, one mentioned that many

faculty keep up with their field "through students that are working on an interesting problem…I think that that's one way, a somewhat efficient way, of trying to stay up on things."

A faculty member's research focus strongly influences their need for information. An epidemiologist spoke of using vital record data and hospitalization data from the Iowa Department of Public Health. A researcher from Community & Behavioral Health whose foci included community health promotion and health disparities and inequity described using publicly-available census data as well as information from the Behavioral Risk Factor Surveillance Survey from the Centers for Disease Control and Prevention, as well as other surveys from the CDC and state health departments. In contrast, a biostatistician described how much of his work as being based on "common knowledge." A statistical method "that's been around for a hundred years, you don't necessarily need to cite what the method is… you don't need to reference that." However, he also mentioned having difficulties keeping track of journal articles as he has transitioned into the work of electronic journals. He previously kept organized files of print articles, but has been frustrated in his attempts to organize electronic copies.

Data

Data – obtaining it, analyzing it, managing it – is at the heart of the research done by Public Health faculty. Their research generally elicits data, and they depend on data from other sources to inform their research. Data was the largest broad concept in both number of codes grouped under it (44) and the number of occurrences of these codes (163). Codes that were grouped together under the broad concept of *Data* included **Data Collection** (35 occurrences), **Data Management** (17 occurrences), and **Data Storage** (16 occurrences). Note that the word "data" is used as a code, denoting information about data in general, but also as a broad concept to group these related codes together. **Data** in bold indicates the code, while *Data* in bold and italicized indicated the broad concept. Table 4 below shows the 44 codes included in this concept.

Table 4: Codes grouped into the broad concept *Data.*

Code	Occurrences	Code	Occurrences
Data Collection	35	Data Conversion	1
Data Management	17	Data Exactness	1
Data Storage	16	Data Incomplete	1
Data Analysis	10	Data Managers	1
Data Elicited from Research	10	Data Ownership	1
Data Visualization Tools	6	Data Provision	1
Big Data	5	Data Sharing Agreements	1
Data	5	Dataset Design	1
Data Sharing	5	Death Certificate Data	1
Data Visualization	5	EMS Data	1
Data Availability	4	Federal Data	1
Datasets	3	Hospital Data	1
Missing Data	3	Incomplete Data	1
Census Data	2	Iowa Cancer Data	1

Data Entry	2	Pharmacy Data	1
Data Extraction	2	Population Data	1
Data from Registries	2	Precollected Data	1
Data Purchasing	2	Primary Data	1
Data Quality	2	Public Dataset	1
Data Usage	2	Quality of Data	1
Databases	2	Relational Databases	1
Data Archiving	1	Sample Size	1

The code **Data Collection** is used 35 times, while the closely related **Data Elicited from Research** is used 10 times. As previously noted, one interviewer used **Data Elicited from Research** to describe data coming from the researcher's own work and used **Data Collection** was used for data coming from other sources. This differentiation was buttressed by one researcher's comment that he assumed that "by elicit, you mean when we actually acquire the data ourselves as with primary data collection." The other interviewer did not make this distinction, using **Data Collection** or other data-related codes for all data. Data sources are also implied in some parts of interviews that were coded with other data-related terms, such as **Big Data**, **Data Availability**, **Data Management**, and **Datasets**.

Researchers described many types of data collection. A researcher whose major focus is prevention and disability described an osteoarthritis study in which the researchers performed MRIs themselves, rather than depending on MRIs taken in other situations. The MRI images were the "data" that they used.

The Iowa Fluoride Study was mentioned as another project for which original data is being gathered. The data comes in "waves" from this longitudinal study as every few years, the 400 study participants, who began as children and are now young adults, receive a dental exam to collect data. Dietary information is also collected in survey form, originally from the participants' parents and now from the participants themselves. According to the researcher, one of the problems that occurs in longitudinal studies of this sort is missing data, since it depends on the long-term cooperation of the participants. He commented that "it's not as much of a problem in the state of Iowa because people tend to be really cooperative." While discussing the problem of missing data, another researcher involved in genetic research discussed how, in his field, records with missing data no longer always have to be discarded. Instead, using "sophisticated analytic techniques" they can sometimes "impute what the value should be for the missing value."

One researcher interviewed was the director of the state-wide cancer registry. He pointed out that the university spends a million dollars per year to collect data on every Iowan diagnosed with cancer, so one of his primary goals is to look for ways to utilize this information. He mentioned that one of the challenges he sees is that although there is a trend to use data from the electronic medical record, "there are a large number of small rural hospitals that are not at the forefront of electronic data collection…they have to educate their people, they have to purchase the hardware, educate them on software, and this is not something that they have a lot of resources to put toward."

The same researcher described a study that crossed the information from the state-wide cancer registry with state's Department of Agriculture database of those licensed to purchase restricted-use pesticides.

Using this information, 60,000 residents were enrolled in a long-term study on the use of pesticides, including the occurrence of cancer.

Re-use of data collected for other purposes was a common theme in the *Data* category. Some data is freely available, while other data, such as that generated by insurance companies, must be purchased. The College of Public Health also collects datasets in the Center for Public Health Statistics. One researcher spoke of the frustrations sometimes engendered by attempting to use data collected elsewhere. He explained that "…how clean the data is in its raw form varies tremendously from one data source to the next, and in some cases it's set up pretty nicely. In other cases you might have to do a lot of merging and cleaning and fixing and so on, generally getting to the original data, the raw data." He also mentioned that sometimes, with large national databases, "you start with a ton of data but you're only going to use parts of it, focusing on… variables of interest. So the first part might be taking this really large dataset and culling from it what it is you need." In other cases, "the investigator has to fill out a form and say this is my study and this is what I want to look and they need to justify … on a need to know basis…You have to say these are the variables that I need and this is the reason why I need them."

A researcher who works with driving studies was able to link crash data with Iowa Department of Corrections data to understand the links between car crashes and citations. He mentioned that administrative data like this "can be really messy but it's really informative and we have also found gold mines of data sitting in unexpected places." For example, schools and workplaces "fill out a lot of forms but then don't collect them as a database, so we are often spending a lot of time going to a partner institution and looking through their piles of records and filling out data from the forms they've collected. This is common among people that are doing community research and I will say that it's very time-consuming." Similarly, another researcher told of spending a year and a half cleaning up administrative data from a major employer, "and what was so interesting is that, once we got it cleaned, actually running the model to give us the answer…took about 8 minutes!"

Researchers working with community-based projects sometimes work with data gathered by their partners, which can lead to similar problems. "One of the things we see is people will have a great big database they're keeping in Excel and they sorted on one column rather than the entire database and so now there's no longer any integrity in the relationships in the data that they have and there's nothing you can do about that."

Changes in gathering research was a common theme, based on the ability of technology to facilitate the gathering of data. For example, two researchers mentioned using computer-assisted telephone interviewing, which facilitates entering responses directly into a database. Another researcher remembered how, in the past, hospitalized patients' vital signs were taken every four hours; now technology makes it possible to gather and store real-time data, "but now the question is what to do with it and how do you take advantage of all of that information? Is there any benefit to looking at a critical care patient in monitoring their temperature every 2 seconds as opposed to every four hours?... So what is a real challenge these days for statisticians is to try to figure out how to analyze these large datasets in such a way that they give added value over and above what we had access to on 20 years ago. I think in general it's much more of a blessing than a curse..."

Data Visualization, along with **Data Visualization Tools**, was recognized as an emerging focus that most researchers felt they needed to learn more about. One researcher pointed out that, for the first time,

the Statistics Department was offering a course on data visualization. Another admitted that his favorite data visualization tool was still PowerPoint, despite its limitations. Several years ago, one researcher from Community and Behavioral Health did a "feedback booklet for the community of what we found in the survey and used mostly just simple pie charts and bar graphs," but says that now "they are talking about using pictographs and things like that." She felt that younger faculty were more adept at visualization, and some that she works with "have used pictographs very effectively in overseas settings to sort of explain to semi-literate populations what the research was showing." Another pointed out that with datasets of millions of records, it is important "to try to come up with ways of presenting the data so that somebody can understand it conveniently without having to look at it in a granular sense… and try to depict what's happening in the forest because there's too many trees."

The interviewers asked the researchers about how they manage and store the data they use. For librarians, **Data Management** includes the organization and indexing of datasets, primarily to make them findable in the future, in addition to storage and security issues. The researchers' answers, though, indicated that they do not think about the organization aspect of data management. One discussed the differences between storage on Windows and Linux computers and using high performance computing clusters. Another discussed using data analysis software. Others discussed the need for secure servers. The amount of storage space needed was frequently mentioned.

Although the use of accessible datasets was a common theme, the idea of **Data Sharing** was not. One researcher felt that data from the driving studies he worked on was not shareable because "you have GPS coordinates, so you could very easily figure out where someone is living." The same researcher was representative of others when he said, "In terms of data, unfortunately, I think that we know that it's important to share data, and we tend to share data on a case-by-case basis with other institutions, once we appropriately anonymize the data and all this sort of stuff. But, to make it more broadly available, there's just so much red tape, or at least perceived red tape, so it's just easier not to do it."

One researcher did talk about using queryable databases to enable searching of data, but he did not mention issues with metadata or terminology used in the queries. Another mentioned that RedCap is useful for surveys, since it builds a database from the collected data. Another mentioned that they have learned to make sure that data is clean and "coded as we expect them to be coded," but this was for their own use, not the possible future use of others.

Publishing

As with all academics, selecting where to publish is a complicated matter. Public Health faculty want their research to be known and to make a difference. The most mentioned venue was peer-reviewed academic journals, but because many of the interviewed researchers also worked in non-academic community-based settings, they also talked about attempts to reach a non-academic audience.

Thirty-two codes, with 104 total occurrences, were grouped under the broad concept of *Publishing*. The most commonly used were **Journals**, **Open Access**, and **Publishing Practices**, with 12 occurrences each. **Publishing** (as a code) has 11 occurrences and **Dissemination Practices** occurred 10 times. Table 4 below shows all codes groups under *Publishing*, including 12 that only occurred once. As with the codes grouped under *Data*, several are redundant, even after standardization.

Table 5: Codes grouped into the broad concept *Publishing.*

Code	Occurrences	Code	Occurrences
Journals	12	Policy Briefs	2
Open Access	12	Predatory Journals	2
Publishing Practices	12	Repositories	2
Publishing	11	Statistical Journals	2
Dissemination Practices	10	Academic Journals	1
Final Research Outputs	4	Associate Editor	1
Nontraditional Publication	4	Citation Management	1
Newsletters	3	Journal Review Process	1
Authorship	2	Magazines	1
Available Journals	2	Output	1
Conference Presentations	2	Policy Development	1
Journal Impact Factor	2	Publication Costs	1
Journal Selection	2	Publication Type	1
Papers	2	Publishing Briefs	1
Peer Review	2	Publishing Papers	1
Peer Reviewed Journals	2	Unofficial Versions of Publication	1

Because many Public Health professionals collaborate with faculty in other fields, they discussed publishing in non-Public Health journals. "Almost all of my research is published in peer-reviewed cancer journals and cancer-related journals, and it runs a wide gamut depending upon what it is that we're looking at." Said another, "I typically publish in the clinical journals of the area that I'm doing the research.   So obviously with trauma, I typically to go the injury and injury control literature, sometime into the critical care literature. With stroke, I typically go to the journal *Stroke* or *Cerebrovascular Diseases*, etc." Another explained that "if we're doing something related to occupational safety in the healthcare industry, we'll target professional nursing journals, because they might be the audience to benefit the most."

Several mentioned publishing outside of health sciences. One seemed to wish that this was not necessary, saying, "I worked with some geographers on an NSF [National Science Foundation] project in Ghana, so some of that goes to geography … journals, which are kind of bizarre." A biostatistician also mentioned sometimes publishing outside of the biomedical field: "[T]here are several Biostats-slash-Stats journals that are targets for myself or for other faculty. There's *Biometrics*, there's *Journal of the American Statistical Association*, there's *Statistics in Medicine*, and there's a variety of others at varying levels of quality and impact, as some people might measure or claim."

Others mentioned specifically targeting Public Health journals. The same researcher who had published in geography journals said, "Most though, are in public health journals, so Health Education and Behavior is one in our field that is pretty well recognized." Others that she mentioned were *American Journal of Public Health*, *Health Promotion*, *Health Education Research*, and *Journal of Public Health Management and Practice*.

The Journal Impact Factor and prestige in general was an obvious component when selecting journals for submission. "We prioritize health-related journals that have high impact factors that are indexed in PubMed." Describing one cohort study concerning the relationship between lifestyle and cancer, a researcher said, "I think we'd probably gotten a dozen to 18 publications out of that particular study, all in fairly well regarded peer-reviewed journals."

There was a divergence of knowledge and opinions on **Open Access** publishing. The same researcher who said that they prioritized high impact journals concluded that statement by saying "More and more, we like open access." Two researchers, though, admitted that they never considered open access journals, and another seemed to equate open access with lower quality: "For us in academia, the difficulty is what I referred to earlier about the importance for a junior faculty being published in high impact journals. That's not the open access venue." Another seemed to be somewhat skeptical of open access, but aware of the trend: "Some of our publications are in open access journals and there are increasingly more of those out there. We usually aim high with most of the work going to a high impact factor journal and they are less likely to be open access, but if they're not accepted, as we move down to get to lower impact factor journals we are more likely to be encountering open access."

The Public Access Policy from the National Institutes of Health (NIH) was also mentioned spontaneously by two researchers in the context of open access publishing, even though, with a one-year embargo allowed, it is not true open access. "When it comes to making publications available through open access, we are required to through NIH, so we're required to make to put those in PubMed Central." Another mentioned it indirectly, discussing the need for articles to have PubMed Central Identification Numbers (PMCIDs) on NIH progress reports. "I get a lot of scrutiny about having PMCIDs. I'm a member the cancer center here, and any time they put out a report or a progress report or an application for funding … they're going to be scrutinized for their PMCIDs."

At least two of the researchers were aware that the University of Iowa Libraries has a method for paying the cost of publishing in open access journals and mentioned it without prompting. One biostatistician related an anecdote about the cost of open access publishing. "I have published in certain open access journals. Actually I've published more in journals in other disciplines because we don't have in statistics quite as many open access journals that have gained the sort of visibility that open access journals have in other disciplines. So for instance, a friend of mine recently submitted a paper in an open access journal. He didn't realize it was open access and when it was accepted he was told it's going to cost $600. He was sort of shocked by that, because he doesn't do interdisciplinary work. He said, 'Well, you know, I didn't realize that was the model' and he didn't want pay it. He corresponded with me about this and I said, 'Well, this is becoming increasingly the norm,' and I talked about how you [the University of Iowa Libraries] have a fund that the investigators can apply to if they want to publish in an open access form to cover that cost. That you know there's a benefit in moving towards that because of these ridiculous institutional subscription rates."

Another mused about the changing landscape of scholarly publishing. "I think the opportunity to publish has increased. I think the ability to publish in *PLOS* journals and other things also has increased. I think the mechanism for publication might be different or the payment for publication might be different. I think in a way, the opportunity to publish has greatly increased, but I also know that journals now try to set a quota of rejection rates and several journals that I know of have 80% to 85% rejection rates. So

their goal is to reject the majority of the articles that they receive. So, there is kind of like this general feeling the article will be published someplace, but it may not be in my journal of choice."

Another talked about the difficulty of publishing in the most appropriate place, while also getting "credit" for the work. "*Health Affairs* and *Journal of Rural Health* hit the audiences that we want to hit. The *Journal of Rural Health's* impact scores are below that of other leading journals like *Health Services Research* in our field. So … if we think we've got a piece of work … that we think could appeal to a journal like *Health Services Research*, we'll put it there because we've already got a policy brief on descriptive stuff that we know is getting out to the general audience that we want to target. Then we can think of little bit more about the traditional academic priority of high impact journals. If we have something that appealed to the *Journal of the American Medical Association* and the *New England Journal of Medicine*, we would certainly go there because in both audiences they're widely read by the policy audience and they are both high impact journals. Most of our work doesn't quite fit what they're looking for. It's not that our work is scientifically less meaningful or less important. It is. It is just a fit issue."

Beyond traditional peer-reviewed journals, though, Public Health researchers disseminate information in a number of other formats. Policy briefs, mentioned above as an alternative to reaching audiences who do not have access to peer-reviewed journals, were mentioned several times. "[W]e publish policy briefs and policy papers from the RuPRI [Rural Policy Research Institute] Center for [Rural Health] Policy Analysis… When we do that, they are posted on our web site. We have an e-mail list. … We have several hundred that get a notification." Another explained, "Work we do on agricultural safety and health, we feel that needs to get to the agricultural community. So then we also spend time creating informational pieces and policy briefs to influence those communities and policy makers." Another mused that "In health management and policy, for example, they have policy briefs, which, you know, are a little different in nature than a typical journal article, but they are still important."

One researcher described repositories for alternate forms of dissemination. "There's a rural repository for toolkits, etc., and we also have some toolkits on our website. One is on an intervention for small restaurants – just putting a table tent that sort of prompts people to think healthy. And then the other was a colleague in Internal Medicine who did a whole thing on food stands at the athletic events and trying to make the food more healthy and … how schools could do that."

A biostatistician described how he decided to publish an algorithm associated with a paper, and the issues that sort of publication could involve. "You'll write up a paper that describes a particular algorithm and you'll think, I'll make it easy for you to use this and so I'm going to write this program in R that will implement this method and then I'll publish it in R. You're sort of responsible for the preparing of the documentation and also debugging it … because if the documentation is really good and understandable people will use your software more. Often you'll also publish a paper where you'll make reference to the software package that accompanies the paper but in the paper itself, you can't talk about how to use the software. You're just talking about the algorithmic method and so the documentation becomes a vital piece of the adoption of the software because if the documentation is really poorly written, people will think, 'Well I don't understand how to use this,' whereas, if it's very nicely laid out, then people will be intrigued by your idea that you published and then they'll read the documentation and they'll say, 'Okay, I think I can figure out how use this package of online data' and then you know they'll use it and that's a real important way of contributing to the profession." He went

on to discuss the fact that, although the paper may be peer-reviewed, the accompanying program probably is not. He also worried that there are no markers like citations to judge whether the program is being used.

Another researcher was hesitant to publish in alternate forms, for fear of not being able to publish the information later in a peer-reviewed journal. "I think the goal is still publication in journals and so I don't want to get the information out that is too widespread in the media." However, he also felt that "we should do more technical reports or publishing under the university's name as opposed to trying to find a journal. That is particularly true for methodological research. The journals are not interested in methodology, per se; they're just interested in results. But we want the methodology out there because it's innovative or different in our paradigm. I think the electronic world is now opening up in terms of what we can do with publication."

Challenges

Like *Data*, *Challenges* was one of the original codes that also became a broader grouping concept. Eighteen codes, with 75 total occurrences, were grouped under the broad concept of *Challenges*, although almost half of these occurrences received the general **Challenges** code. The next most frequently mentioned challenge was **Keeping Up with Trends**, with 13 occurrences. Ten challenges were only mentioned once, although some were reflected in the code **Challenges**. It is important to note that the code **Challenges** was frequently used by both interviewers in conjunction with codes for more specific challenges. For example, in one interview, a single answer was coded with **Challenges**, **Caregiver Overload**, and **Resistant Administrators**.

Table 6: Codes grouped into the broad concept *Challenges.*

Code	Occurrences	Code	Occurrences
Challenges	36	Decisions	1
Keeping Up with Trends	13	Information Overload	1
Time Management	5	Overcommitted	1
Record Keeping	3	Privacy Issues	1
Compliance	2	Publication Access	1
Concerns	2	"Research Nation"	1
Increase Resources	2	Resistant Administrators	1
IRB Approval	2	Scrutinization of Research	1
Caregiver Overload	1	Tenure	1

Challenges seemed to also be used frequently with codes from the *Data* concept, such as **Data Collection** and **Incomplete Data**, reflecting what one researcher called "a real challenge to make sense out of all this data." Some of these data-related challenges were discussed in the section on data, above, including the issue of how much data is too much or, as one faculty member expressed it, "How do we deal with all this big data?" "It's easy to get data electronically, but it's also harder to maintain control of the data or security of the data."

Another data-related concern was missing data and how to deal with it. "Sometimes in the analysis you delete those individuals but increasingly you have other information on them, but you just don't have it for this particular variable. So people are trying to look at other people very similar to this and impute what the value should be for the missing value. So you get into some sophisticated, analytic techniques." Another researcher declare that "I think public health data is inherently unclean! What are you going to do? … I think a lot of it is a challenge of trying to weight the data and then do some data imputation because of missing data and that kind of thing."

Archiving data was also discussed by several faculty members, particularly when the data might be useful in the future. To the researchers, this appeared to be different from data management, although to librarians, data archiving is part of data management. "…I don't think we've done very well at archiving data, maintaining data from existing studies or past studies, so that they can be utilized in the future. Part of that is a software issue, part of that is also a storage issue. I think that to me those are big challenges for future." Another researcher pointed out that "…you know you might not necessarily want to save all of the original data but you need to save enough key information from each simulated data set so that you can analyze results appropriately."

In many studies, before data can be gathered, study participants must be recruited and then retained, which was one researcher's major challenge, although she said that "I think we've done better with retention because it's mostly intervention research, so people had a benefit of staying and they were receiving a service, etc. But certainly recruitment has been a real challenge." Another spoke of challenges in contacting the individuals he needs in his research. "Trying to recruit direct care providers [for individuals with dementia] has been a challenge, because naturally you try to access them through the administrators, but when I go and ask them about potentially their employees participating, I usually get a no answer. So that has been the biggest challenge." Another researcher reflected on changes that affect recruitment. "I think it's just inherent in what we do, in a busier world. …You know, cell phones make our life easier but it makes recruitment hell."

Change was frequently seen as a challenge, including with funding. Because most public health research and public health programs are government-funded, the future of funding was a frequent concern. "The funding seems to be shrinking.   So that is a big concern and of course we are concerned about research funding, but I've been talking with my community partners over the past few days, who work for the states in health department and aging departments. They are very concerned about budget cuts and what programs might have to go. So I think that is the biggest concern that I have." Said another, "There's going to be a real focus on personal medicine and some of us believe that prevention might go a lot further than personalized medicine. So, you know, I think that might impact what funding is available and we've already seen it [a change in funding] in practice."

Change at more mundane levels also affected the faculty members. "We have students that we use as graduate research assistants and if we have a study that goes five years and these are masters students that are only here for two years, then you can tell that there's going to be two or three students that are going to work on a number of these papers …, so we need to train them to make sure that they document are steps clearly so that is easily passed on to the next person."

Concern about the future presented a challenge to several researchers who worried about how their work would be perceived. One researcher expressed concern over future **Scrutinization of Research** and how it might reflect on his legacy in the field. "I typically don't do real high risk research but nonetheless

thinking about the fact that someone will come along and say, 'Well, you didn't do ethical research when you were conducting it'—not that it wasn't done according to rules at the time, but they are applying current rules to something that was done 20 or 30 years ago, which is inappropriate …people want me to answer for something that is current but, to apply it to something it was done years in the past."

"I'm not dealing with the same people that I did 10 years or 15 years ago. Or the same operating methods or the same style in terms of management and, in a way, different constraints, etc. Part of it may be the fact that they're trying to develop best practices or develop a definable strategy, but basically I think it is new medicine forgetting past history of how we used to do things or how we could do things."

A researcher worried about the future of the U.S. as a leading **Research Nation**. "I also have concerns about how long the U.S. is going to be the top level leading nation on the planet ... Particularly with the trillion dollar debt which is growing, so all these things make me concerned about research in a sense, because research is typically done by countries that have money to do it."

Conclusion

The Ithaka study attempts to take a broad view of the research practices and related information needs of academic public health faculty. The interviews with eight faculty members in five different divisions of the College of Public Health were interesting and informative in that regard and will prove useful to the department liaison in the future, since he has greater insights into the College's research priorities.

The librarian-interviewers were also able to discern areas where they might be able to be of assistance. For example, although most funders now require rudimentary data management plans, data management is a significant challenge. It appeared that most faculty were not concerned with how data is described and documented and therefore made findable and usable by others, either inside their department or more widely. This is an area where the University of Iowa Libraries' Data Services Manager could offer advice.

The discussions also made the interviewers more aware of the numerous research centers in the College of Public Health and the many policy briefs and other publications these centers produce. Many of these are posted in various locations on the College's website. The librarians involved in the interviews hope to explore methods in which these groups of papers could be interfaced with the Libraries' institutional repository, Iowa Research Online, which would benefit both the research centers and the institutional repository.

Although the interviewers felt the discussions were worthwhile – speaking to faculty about their research is always enlightening and usually enjoyable – the lack of information gleaned on use of library resources was disappointing. None of the interview questions focused specifically on the library as an information resource or on the researchers' use of library owned or licensed resources. When asked about what kinds of information, beyond their own data, they required in their research, the faculty members tended to focus on their need for government datasets. Most of the researchers did not mention accessing information through the library, although one researcher discussed her frustration with searching MEDLINE and Web of Science and described her research group as "big, big library

users." Another researcher said his group relies "quite a bit on the institution memberships in a lot of the journals." In general, journal articles and book chapters were discussed as works they produced, not that they used.

The information needs of public health workers such as those working in state and local public health departments and clinics have been studied in sufficient numbers to produce a systematic review on the topic, [1] but there is only one article, [2] more than ten years old, that focuses solely upon the needs of public health academic faculty and researchers. In that article, the use of library resources at a single college of public health was studied. The interviewers had hoped that this project would give insight as to how researchers at our own university used library-provided resources, and how this compares to use at other institutions. The fact that this information was not gleaned from the interviews may have been a consequence of inexperienced interviewers who failed to follow up with sufficiently probing questions, but may have also been caused by the wording of the questions themselves.

In spite of this disappointment, the interviewers now have more awareness of the research activities of the College of Public Health and will be able to build on this knowledge to enhance our services to the faculty and students.

References

1. Barr-Walker J. Evidence-based information needs of public health workers: a systematized review. Journal of the Medical Library Association. 2017; 105(1):69-79.
2. Wallis LC. Information-seeking behavior of faculty in one school of public health. Journal of the Medical Library Association. 2006; 94(4):442-6.