2021 NCSBN Scientific Symposium - Keynote: Open Science, Public Accountability: NLM Helps Nurse Scholars Shape Public Discourse Video Transcript

Event
2021 NCSBN Scientific Symposium

More info: ncsbn.org/15185.htm

Presenter
Patricia Flatley Brennan, PhD, RN, Director, National Library of Medicine, National Institutes of Health, U.S. Department of Health and Human Services

- [Maryann] Hello and welcome to NCSBN's Scientific Symposium. I'm Maryann Alexander, Chief Officer of Nursing Regulation at NCSBN. The studies being presented today were conducted both by NCSBN's internal research staff, and external scientists funded through NCSBN's Center for Regulatory Excellence grant program.

The grant program awards over $1 million annually for studies that advance the science of nursing regulation. The data from these studies are used to make nursing regulatory and policy decisions, develop national guidelines for nursing practice and education, develop important changes to nurse practice acts and regulations.

We want to thank all the researchers and grantees for their important contributions to nursing science and a special thank you for those presenting at this year's Scientific Symposium. One of NCSBN's researchers was Jennifer Hayden. She was a young talented scientist that was the principal investigator for NCSBN's National Simulation Study.

This study changed the way simulation is incorporated into the pre-licensure nursing curriculum around the world. Unfortunately, almost upon completion of the study, Jennifer passed away from breast cancer. Because of her important contributions to NCSBN and to nursing science and policy, we have named our Scientific Symposium keynote address in her memory.

And now it is my pleasure to introduce to you the Jennifer Hayden keynote speaker. This year, we are honored to have with us Dr. Patricia Flatley Brennan. Dr. Brennan is the Director of the National Library of Medicine at NIH. She is the first woman and the first nurse to be named to this prestigious position.

Prior to this she was the Lillian L. Moehlman Bascom Professor at the School of Nursing and College of Engineering at the University of Wisconsin in Madison. Currently, Dr.

Brennan is working on the delivery of health information in new ways that are centered around the patient experience. Her laboratory develops interactive virtual reality experiences to better characterize patients with complex chronic conditions such as diabetes and heart failure, and a context in which they live. These simulations have several purposes, including enabling patients to rehearse problem-solving behaviors, to help improve their health outcomes, and disease management skills, inspiring design of innovative home care technologies, and improving understanding of the sensory behavior and cognitive processes that shape self-care.

Dr. Brennan was on the faculty at Case Western for 10 years and for 20 years at the University of Wisconsin in Madison. She also is proudly the mother of a 20-year-old son named Connor. Thank you, and please join me in welcoming her to NCSBN's Scientific Symposium.

- [Dr. Brennan] Good morning. I'm delighted to be here and quite honored to be the Jennifer Hayden speaker today.

I'm bringing you ideas about open science and the way nursing and the National Library of Medicine can work together. We will be, at this next 40 minutes, having presentations and... Excuse me, I'm not seeing my slides. We'll be having presentations and conversations.

This is an audience participation activity and we'll be here together with you to discuss the different ways that the NLM can help your scholars shape the public discourse. There are three objectives for the session today. First, to recognize the role of the National Library of Medicine in supporting the scientific responses to health challenges.

Second, to critically appraise the contribution of the NLM's offerings of bibliographic and full-text literature databases, biomedical databases, and repositories such as clinicaltrials.gov in support of nursing science. And, finally, to devise pathways for public discourse that enhance the impact of one's science. Now, I recognize this is quite a large audience and quite a mixed audience so you may have different goals for what you would like to get out of my session.

We are going to be having audience engagement throughout the entire session and we'll begin now with the first poll. If you go to your screen, to the right-hand side, you should see a poll that lists these three objectives. Please select the one that is most important to you. So, actually, interestingly enough, we're sort of evenly divided. Lots of interest in what the National Library of Medicine does, about 35% of the responders.

Not as much interest in recognizing the role the library has played in supporting science but a lot of interest in devising pathways for enhancing the public health contribution of one science. To begin understanding the role of the National Library of Medicine and what our contributions are to nurses cholars and how they can shape the public discourse, please watch this short video.

♪ [music] ♪ Over the next 40 minutes, you're going to learn a lot about the National Library of Medicine. And you're probably going to come away with the idea that it's not your mother's library anymore.

We're very excited to be part of the National Institutes of Health, which we joined in 1966. However, our library is 180 years old. Right now, in the 21st century, we focus on critical areas, critical infrastructure for knowledge and policy. We facilitate open access to the literature.

We have resources such as PubMed Central, our full-text literature repository, the CORD-19 recollection which has over 120,000 COVID-specific articles that had been made open to the public for machine learning as well as for general perusing during this terrible pandemic, and PubMed, the bibliographic citation database that we have.

We also conduct and support research in computational biology and computational health sciences. And we implement and establish training programs. We have training programs around the country, 16 programs in pre and post-doctoral training in biomedical informatics. We host hundreds of training programs throughout the year for clinicians, patients, and librarians to better understand data science and the resources of the National Library of Medicine.

And we focus on informing policy. Now, as a federal body, we can't make policy but we do provide the educational and informational resources to shape policy around open data, research integrity, information access, and research accountability. Our focus today is on open science and public accountability.

And the phrase that comes up quite often within open science concepts is the idea of open access. Now nurses, particularly nurses in practice who graduated from their programs that have been richly supported by libraries suddenly find that they can't always get access to the journals they're used to getting access to. The concept of open access means bringing the literature in the open to all who may need it.

We're going to have a brief poll now to talk a little bit more about specifically what does it mean for the library to open access to literature and data? There's a new poll that's upped just now. Please respond. I'll try to give it a little more time this time so that everyone gets a chance to respond.

What is open access? Not everyone will know the answer to this. Don't be concerned. We want to make sure we start off at the same point, though.

There's a new poll that's upped just now. Please respond. So please respond. Open access, 42% of you got this correct.

It's a set of principles and a range of practices through which research outputs including journal articles, but not only journal articles, are distributed online, free of cost, or other access barriers. Now often, open access is thought of as the publishers making a decision to remove the paywalls. And in fact, very often, that is a critical piece of open access, that is publisher's charge through subscription or through other mechanisms to get access to journals.

But generally, from the perspective of the library, the open access is a partnership that allows for access to information generated by researchers supported by research activities and communicated through many mechanisms, including publications, journal repositories, and journal resources.

The National Library of Medicine is building the 21st Century Collection. The 21st Century Collection has three key responsibilities. First and foremost, we preserve. We preserve materials that are available for the centuries to use. You might be surprised to know that in our library, we house 10th-century manuscripts from China that have helped us understand some of the basics of neuropathic and plant-based interventions.

But increasingly, libraries are becoming electronic. And so we connect to other resources. We connect to the publishers, we connect to different sources of information, data repositories. In the future, discovery in the moment will become important. What is the answer to the question that you haven't yet thought of?

Many of our resources are now connected because we create elaborate interconnections through search strategies, but in the future, we have a need to think about and anticipate the information needs as they unfold. Now, if you look across the bottom of the screen, our primary substrate for libraries has always been literature. But also I want you to think about that globe in the center, an interconnected set of literature, data, other products of research, code, pipelines that may be of value to document and show the rigor and reproducibility of our work.

And of course, increasingly, data are important. Now, the National Library of Medicine has two key resources that are most familiar to the general public and to nurses. First is PubMed, a search engine that accesses the MEDLINE databases of reference, citations to articles in PubMed Central and some other related materials. We have over 30 million citations in PubMed.

PubMed Central is a free digital repository that archives publicly accessible full scholarly articles and serves as the public archive for the NIH Public Policy. and for many other federal institutions, including, for example, the Department of Agriculture. We have these resources available 24 hours a day, 7 days a week. But during this period of time, with the tremendous public health emergency we have going on, we established the COVID-19 resource and began making available from first with a partnership of 50 publishers, over 95,000 articles that were made accessible and machine-accessible.

This allowed for increasing the discovery process, identifying relationships between various medic drugs that are existing in use, and new possible targets for the COVID pandemic. This partnership required that we work closely with the publishers and the Office of Science, Technology, and Policy to remove the legal and financial barriers to accessing this information.

I'll be glad to talk about that more in the question and answer session if you have some interest here. In addition to PubMed Central, we're partnering with the Allen Institute to build the CORD-19 data set. This has provided a challenge base for other investigators, for individuals with clever ideas, to be able to mine and extract this information.

The concept of preprints, that is publishing a manuscript prior to a journal review, is getting greater hold in many disciplines, including the Health Sciences. The NIH began accepting preprints as a credible evidence of progress towards research goals about two years ago.

Starting in the spring of this year, in June, the NLM launched the preprint pilot, where we now are making preprints discoverable through PubMed Central and through PubMed. We're focusing at this point only on the COVID-19 articles, related articles. But what this is allowing is the literature now, before it has gone through peer review, can be quickly accessible to researchers.

Of course, this requires that we be sure that people understand our brand, and make sure that they don't confuse a preprint with an actual archive article that's been reviewed. So we use a lot of banner headings and electronic indicators to identify and differentiate what is a preprint from what is an actual article. We are only doing this for NIH-funded research because we know we have a trusted brand there.

Improving access to the research literature is a critical, critical requirement of the library. As I indicated earlier, we have over 30 million citations. Two and a half million users come every day to use these resources, and conduct millions and millions of searches throughout the month.

Now, what this means if we have 30 million resources, 30 million searches, we have too much literature for you to read all at once. And we know also that an individual who conducts a PubMed search often gets pages and pages of citations. We've been working, over the last couple of years, to integrate artificial intelligence solutions to make sure it's possible to get access to the most relevant literature.

We've developed something that we call the best match process using a learning-to-rank artificial intelligence scheme. So when your query comes in at point number one, we translate it by mapping it to known vocabularies and terminology, and extract from our 30 million citations what might be the most likely top hits for you.

But this still may be too many pages for you, to many papers for you to review efficiently. So then we apply this artificial intelligence learning-to-rank algorithm, which uses thousands of features. How often has this article been cited? What else has been cited along with this article? How recent is this article? And then we return to the individual, return to the searcher what we call the best match. This differentiates from our previous approach where we used to provide our resources in a reverse chronological order, that is most recent first, and you can still get that.

But what we're finding is by using our AI mechanisms, we're able to get information into people's hands faster. What it looks like on the screen is you see in front of you. In the right-hand side, in the green box, you know, you can select either Best match or Most recent. And when you select the Best Match, a pop-up box appears that gives you... and identifies for you the best citations you could possibly get.

We're continuing to improve this both by understanding how much click-through, how much do people actually read the citations we provide, and also by getting responses and feedback from our stakeholders around the world. This best match algorithm, according to our users, does produce better PubMed searches. But importantly, what it means for our researcher is that your work now is not going to be buried on page 30.

It may be on page 1 of the citations. And since over 80% of the people who use the PubMed resource only go to the first page and they don't go beyond that, it's great to get things pushed up front quicker. We envision a world where the PubMed Clinical Queries provide quick, customized access to facilitate discovery of the PubMed resources, getting them available as quickly as possible to individuals.

We run these filters by topic, research topic categories but we're also able to help people create individual clinical queries for their own clinical populations. This does help us to accelerate discovery and opens up the literature more quickly. Opening up the literature is not only making the literature available but making it accessible in ways that people can consume it.

A good deal of our work though focuses on research data sets and making research data sets reusable and discoverable around the world. On the screen in front of you, you see a number of different important research data sets such as the Kids First study from Pennsylvania or the All of Us, the million participants of the… participants to understand health and everyday living, Cancer Moonshot, the Framingham Study, and the NCBI.

That's our National Center for Biotechnology Information. Finding ways to connect these data sets and make them available by a query not only accelerates research, but it also provides a wonderful platform for training for our students. And for those of you who are looking to shape public policy, being able to access data directly provides a strengthening to the base of the policies you're trying to shape.

Most recently, we've been looking at how to better support the use of model organisms in research. Model organisms are important because they provide a way to better understand a particular physiologic oriented process that's relevant to humans, but can be studied in another species. In our case, most of our model organisms provide part replicas of humans like zebrafish, or rats or mice.

Some of you may have even studied some of these in school. But to build the knowledge base that connects them together requires that we construct the ecosystem that you see on the left-hand side here. So on the upper part of that left-hand blue document, you see various organisms, the sequences, and the anatomy, and images about those organisms. In the two circles, you see the data related to those organisms, transcriptomes, genes, orthologs, as well as the tools to interrogate the data.

The BLAST search was just a broad analysis search or genome annotation. The role of the National Library of Medicine is to bring those two circles together with the communities that want to study them. And we also work very closely with those communities to annotate and to provide the critical description of what gene structure appears in multiple organisms, for example, or doesn't have the same function in multiple organisms.

So, our library is designed not only for human readers but also for machine processing of large and vast amounts of data. Our goals are to provide a central portal, the ability to work in the cloud, the ability to have large-scale compute properties that don't require a single institution to hold a number of resources, to include shared tools and scalable analysis that provides the infrastructure for data science.

Now, nursing is just beginning to really exploit these resources. And I'm looking forward to making them even more useful to the nursing community. The National Library of Medicine, in addition to providing a wide range of resources to communicate about and support the conduct of research, also conducts research itself. Let me pause here and take you to another poll and ask you to… Did you know know that the National Library of Medicine has a… funds both internal research in our campus at Bethesda, as well as extramural research?

Most of you did not know that the National Library of Medicine funded research. But now I'm delighted to tell you that we spend over $60 million around the country every year creating the kind of research infrastructure necessary to support patient care, as well as necessary to support effective information management use.

On the screen in front of you, you see our broad-based research portfolio. Our research portfolio focuses on understanding the machine learning tools that can be used in a variety of resources but particularly applied to learning from electronic health records. We support information standards and discovery.

We do a lot of our work on image processing and creating a way to use machine learning to interpret images. Our focus here is not so much to cure cancer but to build better tools and make the tools better available for the analysts who are working to cure cancer. We focus on natural language processing a lot, extracting meaning from text but also on statistical analysis and our biologically-driven research makes use of our genomic databases to be sure that we can better understand and have access to the secrets that are hidden within genes.

A lot of our work does focus on the electronic health record as a source of information and engagement for understanding. What we're trying to envision is a future of understanding how an individual from the level of their genetic structure to their functional everyday being, to their image, sorry, their anatomy and physiology to bring this together in a way for us to better understand and therefore, as a society, better improve the health for all.

Some of the work that we do focus very specifically on tools that could be of use to nurses. A project in particular is a project called the THYME Project. This is conducted in Boston by Guergana Savova's group. And basically what she's learning to do is to take a broad range of clinical information from the clinical text and begin to provide sequence narratives.

Now, she's using natural language processing to create patient timelines and then annotate these within a document so that a clinician who's looking across the broad range of a patient's care challenges can see what has happened in what sequence. Now, what the National Library of Medicine does uniquely that other institutes in NIH don't do is we try to develop reusable tools.

So as Dr. Savova's group is developing an understanding of what constitutes a disorder and how do we understand the clinical narrative over time, she's also making those tools available for other investigators and for other health care systems to be able to make use and reuse her information. We are quite interested in better support for clinical trials. As you know, the COVID-19 pandemic recognized and relied on clinical studies engaging patients very quickly.

And yet in order to conduct clinical trials, we need to have strong and good record systems so that the trial, that particular protocol, the inclusion criteria are all available. We have a fantastic resource called ClinicalTrials.gov, which is a place where all research, clinical trial research that is designed either to be conducted in support of an FDA application or is supported by the NIH must be registered.

Excuse me, the results can also be reported there. This ClinicalTrials.gov repository actually allows for clinical trials as well as patients to better find out what studies might be available. What's going on in

their area that they might find relevant to someone that they care about or help them in their own management of a disease.

Since 2019, we also report, within one year of completion, the final outcomes of every clinical trial. This is a really important public accountability responsibility that the NLM takes on. We've learned that fully one-third of all clinical trials funded by NIH never make it into a clinical record, never make it into an archival publication, never even get to a preprint.

So by having a structured way to report the outcomes of clinical trials, NLM is supporting rigor and reproducibility in clinical trials research. We're also trying to find ways to match patients to trials better. Where, in this case, we take the eligibility criteria that is in the narrative description in the ClinicalTrials.gov application and build tools that allow an investigator to scan through clinical records safely and protecting patient privacy to identify people that might be good candidates to participate in a research program.

The National Library of Medicine then supports its own research, focusing on reusable methods and better use of data including clinical data and supports the research of others by providing registries and repositories such as ClinicalTrials.gov. I now want to talk a little bit about what we do for training before we start going into public policy. On your polls now, please turn to our question about training.

The NLM recognizes that data science is new to most people and everyone's going to need some kind of training. So what type of training do you think is your preferred means of learning? Well, I see a strong interest in webinars, short, informative focused, and flexible. Some preference for Zoom and no one wants injestible biochips.

That's a pretty bizarre idea. But actually, I had a colleague about 20 years ago advocating that we should somehow find a way to encapsulate knowledge into capsules and then swallow it. Never saw that one as becoming a great use for us, particularly for professional development. But the National Library of Medicine is deeply involved in training across the career lifespan.

So let's talk about some of the things that we do for expanding and enhancing research training for biomedical informatics and data science. I see a strong interest in webinars, short, informative, focused, and flexible. Some preference for Zoom, sorry about that. The training that we do at the National Library of Medicine focuses on skills training for biomedical informatics and data science.

We provide training to undergraduates, to doctoral students, and to postdoctoral students. Nurses are eligible to participate in our training programs, and many of our training programs exist in schools and universities that have strong partnerships with schools of nursing. These training programs provide support and a stipend for an individual to study for several years. The important goal in the beginning of our training program is to begin...to come with an idea, come with a thought, come with a question that can be answered by developing better methods and better use of data and information.

We also provide training across society. We recognize that laboratory scientists, patients, and clinicians all are being faced with this data science revolution. And it's really quite a step away from what we've learned in the developing evidence-based practice or some of what we learned in our research programs

in school. So we are focusing a lot on the use of webinars, particularly webinars that make the library's resources more useful to people around the world.

But we don't stop with professionals. We also are trying to reach into the next generation of lay people, of consumers, and librarians to make sure that the data revolution that's coming is available to everyone. In the screen in front of you, you see the map of the United States and those star points indicate the regions of the network of theNational Library of Medicine.

The National Library of Medicine network is 8,000 points of presence around the country that are driven to best understanding to best development of the library's resources in the community. So we have close connections with communities, we have individuals that can help train or help make accessible our own resources.

We use this now, we've used this several times in the All of Us program. We provide the community engagement structure to make sure that there was somebody present in every community to help answer questions in the...under the seal, sorry, the Heal Initiative that's helped to end addictions long term. Our network of the National Libraries of Medicine provided a place for disseminating information, and also for getting information to the local clinicians and researchers about new initiatives and new strategies.

Under the COVID-19 pandemic, our network once again has come into play. Our network includes everything from hospital libraries to academic health science centers, to public libraries in the community. And each of those serves slightly different people, but make sure the resources, the National Library of Medicine, come forward and can be present. Now, we're starting to assist with information about vaccines and helping to improve vaccine acceptance across the country.

Our training programs, sometimes are degree-granting, sometimes are certificate-producing, and sometimes are information-generating for individuals, all are open and freely available around the country, and are freely available almost online everywhere in the world. I now want to turn or my last set of comments in our last 25 minutes or so to talk about the key concepts for this meeting, which is how do we take all these resources and really begin to inform policy?

I have two polling questions for you right now because I want to better understand the policies that are most important to you. The first question I want you to answer is, what is the most critical public policy that your expertise could help inform and shape? Thanks and thanks to whoever's managing the poll.

I appreciate that.

So, the major public policy concern that was identified by this group as critical is promoting patient safety in hospitals. Almost 60% of you identified that as your key public policy. There are others on here. Some are going to write into the chat with the public policy issues that they had... that they're interested in.

There's about 18% of the people that were going to write something themselves. There is the other policy issues that generated interest though were reducing maternal mortality. About 10% of you are interested in that. Ensuring food security for all. About 5% are interested in that. And preventing elder

abuse, about 8% are interested in that. Now, I'm going to take you to this next poll which is a question about nursing-related policies.

And I hope you understand the distinction between public policy and nursing-related policies. Nurses must participate in shaping the policies of health for all, but we also must participate in the policies that are relevant to our discipline.

So, on the screen in front of you, and in the poll rather, you have a set of four options or opportunity to write your own into the chat. Well, now, I see a rousing 38% think we need to focus on workforce development. But 35% of you are interested in practicing at the top of our licenses.

Only 5, excuse me, 3% are interested in pay equity and about 5% of you will write something into the chat, which we'll take a look at in a few minutes. There's many, many, many public policies that nurses need to participate in, and many ways that we will be creating for the future the kind of policies that are useful to bring the health goals that we want to have for our society.

Today's focus is largely on data science and open science. That is making sure that the large amount of data that we have available is useful and used for the purposes of health and health science. When we move towards a model of open science, we need to recognize that there are several policy issues that have to come into play and interconnect.

And perhaps none is more important than the idea of patient privacy. How do we make sure we can leverage our understanding of data, whether it's knowing where you've been walking with your cell phone or knowing your genetic code, in a way that advances society without exploiting individuals. Nurses must participate in these discussions. These discussions are happening at all levels, from the local to the state, to the county levels, and they are providing a way for individuals to really explore the challenge, essentially the trade-off between privacy and science.

Now, we can make the future that we need to have happen if nurses are engaged in understanding public policy and shaping it. So I want to focus my comments now on what are nurse scholars' special roles in using data-driven science to shape policy. We can do things that no other health profession can do. So I'd like to consider your…to think about our policy activities as being grounded in our base discipline.

And first and foremost, recognize the dignity of the source. That is respecting the dignity of the individual, the person, the family, the community that's providing us with information should also be benefiting from the provision of that information, including being able to be assured of privacy and confidentiality.

Next, it's critically important to have principled approaches to discovery. Our nursing science traditions are strong and theory-driven, and grounded in the ability to better understand phenomenon because we use theory. This brings us to a way of understanding patients and the patient experience in novel ways.

So principled approaches discovery and use of theory to guide interpretation are critical but they allow us to take special attention to the phenomenon of concern to nursing: the diagnosis and treatment of human responses. No other discipline focuses on the diagnosis and treatment of human responses.

So when that knowledge is needed to shape policy, we must be the ones that bring it forward and make sure it's incorporated. We must remember that we should ask and answer questions that are germane to the discipline and to the people that we serve. We begin by thinking about how do we enhance the impact of our science.

Being a great researcher writing terrific papers is only part of the story. Creating pathways for public discourse, that's what I want to focus on for the last couple of minutes of my talk. But you must begin always with this level of rigor and reproducibility. Good science is the foundation of good policies, and it's our [inaduble].

Secondly, a step must begin before we begin anything and that is to build your network. Build your networking conversation, not just in conversation with other researchers, but with the people who your research is designed to help and also with the people who helped to shape and manage the policies that are relevant to the care approaches that you want to have happened.

Building a network is a career lifetime commitment. It doesn't happen project by project. And being known, sharing your willingness to work on policy, sharing your concerns about policies with other policymakers is a way of framing this network. The next step is to know what you know and to know what you don't know.

But often, as we shape public policy, we have a focus on a thematic concern. But we need to be sure we also bring together those who understand the analytics that are used to highlight that thematic concern, the statistics, or the data science that underlies an activity. It is not solely the responsibility of nurses to know everything, but to know that everything that's critical to their policy concept is in fact made present, made available to the group.

Amplifying your story comes in part by preserving trust and provenance in a highly distributed information environment. We well recognize that we're deeply suffering from an information-disinformation tension in our society right now.

But for nursing, to get their materials explained and available to the public and thoughtfully entered into the public policy debate really requires that we use strategies that amplify trust. and provenances the consistency of a message and the source of the message in a highly distributed information environment.

When we make use of Twitter or other kinds of public social media, we have to remember that our job is to leverage these in a way that is trustable and amplifying a message, not adding to the confusion of the conversation. We need to think about how to leverage and how to make good use of these emerging technologies. How could artificial intelligence or augmented reality help to bring forward an illustration of the policy you're trying to shape or the consequences of that policy?

Considering how different ways of communicating, providing multiple channels of interaction, will allow your research to be presented to a number of different audiences. The opportunities to engage with the public have never been greater,. The chance for nurses to shape, to build their networks, to amplify their stories will come from effective use of emerging technologies.

And please don't forget that storytelling is critical in policy development. The plural of anecdotes isn't data. So just having lots and lots of stories doesn't replace the need for data-driven sciences to shape policies but the stories are what sell the science. Stories and storytelling of the impact, the consequences, the opportunity, the ability, the unique perspective that your research brings to a particular problem takes it a long way to helping others who have the ability to shape and decide on the policies, which policies should be attended to and how they should be enacted.

And, finally, please think carefully about transporting. How do you transform science into action? Sometimes, it's through clinical care. Sometimes, it's through the ability to engage with individuals and develop a story that they can then take forward. Sometimes, it's the changing of a product or developing of a whole new model of a community engagement.

There's lots of ways that we transport our science into action and nursing creativity can help us go a long way in addressing the critical policies concerns that can be answered and addressed by the knowledge that we build uniquely as nursing around society. I want to close my remarks by reminding you that the National Library of Medicine has, for almost 200 years, served as a trusted source of health information.

That trusted source of health information benefits from nursing and nursing knowledge and should support and develop nursing and nursing knowledge as we go forward. I look forward to continuing to work with you in any way that I can. You can reach me on email or on Twitter or every Wednesday morning, the blog Musings from the Mezzanine comes out.

I try to showcase nursing stories periodically through the year. If one of you has an idea that you'd like to include in a Musings blog post, please reach out and get in touch with me. I often have guest bloggers. Thanks very much for your time. Thank you for your patience with the technologies. Thanks to the tech people who've been very supportive through this session and I believe I've left us about 10 minutes for some questions and conversation.

- Dr. Brennan, thank you very much. That was an excellent presentation and extremely informative and enlightening.

I'd like to open it up now for questions for Dr. Brennan. Please put them in the Q&A section on your website. Dr. Brennan, the first question I have is from Carrie Downing. She asks, "How many clinical trials do not make it to publication?" She says she knows you mentioned it but she missed the number.

- Carrie, this is a number I'm not very proud to be restating again. But from the work of Deborah Zarin, Z-A-R-I-N, over the last five years, we think that it's about one-third of all clinical trials that were funded by the NIH never appear on publication. Now, there are other clinical trials that are not funded by the NIH that don't appear, don't ever make it to publication.

And frankly, we also have challenges with clinical trials that are conducted internationally. Sometimes, the challenge is because there are… there's a difference in the way studies are conducted, there's culture around sampling plans and an analytic plan. Sometimes, there's no results and journals are actually much more likely to publish positive results or results that at least are definitive, rather than a study that comes to the end and says, "Well, we don't really know that this made a difference ."

One of the critical pieces about communication about studies is knowing what failed is as important as knowing what worked. And that kind of information can be found by looking at the ClinicalTrials.gov repository to look at our results.

- Thank you. I have another question for you from Kathy Scott. She asks, "Could you say more about open science publications and the responsibility of the author versus the publisher?"

- I can say to…from the perspective of the director of the National Library of Medicine but I'm not a publisher so I can't speak specifically for them. When you say open science publications, I believe you might be referring to open access publication. In the PubMed Central repository we have something referred to as the open access subset. These are journals where the journals themselves are completely open.

Now, in most open access journals, that is, the journal makes all of its publications, all of its articles available without a paywall and available electronically, in most open access publications, there is a process called author use fees or author publication fees. And this is a fee that the author pays, as the publication is being developed, to support the cost of managing the journal.

So in many cases, author fees or other publication fees are required to support an open-access journal. The National Institutes of Health does allow investigators to include in their grant budget fees that would need to be paid for open access journal. So the fee can be charged to a grant budget and that does help support the ability to publish in a lot of different outlets.

We encourage authors to use open access journals where possible. The National Institutes of Health has a wonderful set of recommendations of how to pick a good journal to publish in. And this is a way to help ensure that your materials reach the most people. Now, publishing in a what we would consider a traditional journal isn't necessarily a bad thing. And the journals are beginning to become more flexible in their approaches.

So even the journal of my field, the Journal of the American Medical Informatics Association has an ability to unlock journal articles. That is we can open journal articles, remove the paywall for everyone. This then allows information to be distributed much more quickly. And just as a reminder, the NIH does have a policy, that within one year, any articles describing research supported by an NIH funding must be completely open access.

That is the journal can hold an embargo for one year. But after one year the article has to be freely available to the public. In the last year or so, we've been seeing new initiatives come through the Congress that have required that if we're going to receive special funds for a certain kind of research at NIH, we must make the journal…all the articles funded under those special initiatives really accessible immediately at the point of publication.

And that includes the 21st Century Cures, which is funding a lot of cancer research right now, and the Heal Initiative, helping to end addictions long term. In both of these cases, the Congress has recognized the urgency of getting information out to the public really requires that the information be available as quickly as possible.

- Great. Thank you. And I have another question from Supa Tula who asks, "Do you see the term 'evidence-based research' being replaced in light of the expanding field of data science?"

- This is a really good question, and I'm really glad you brought it up because we worked really hard in nursing to talk about evidence-based practice and now we're kind of changing the game a little bit. And people want to know, well, what's really different? What's new here? Many of the data-driven studies, the data science studies that we are involved with do not provide the same kind of evidence that traditional clinical trials provide.

So there will always be a place for the evidence generated through clinical trials. However, more importantly, is for individuals to consider what constitutes evidence. And we see with data science a broadening of what constitutes evidence. Many data science studies are hypothesis-generating as opposed to hypothesis-resolving.

So a data science study might identify patterns that had not previously been identified, inspect those patterns, and provide ways to then subsequently do testing different kinds of evaluation. I also see a trend towards expanding from clinical trials to observational studies, and practical clinical trials as another kind of evidence.

Essentially, the concept of clinical trials is a highly controlled process where only under certain conditions is a very specified intervention evaluated against other interventions or against a usual care treatment. Yet sometimes, the constraints of a clinical trial are too constraining and they don't depict the reality of everyday living.

In my own work, what I've noticed is that the way people organize health information in their house actually varies and affects their ability to remember to take their medications. So, when I work with people who are developing homecare clinical trials, I try to encourage them to take a map of the house, take a look at what else is going on there. Now, in your future, what I hope we'll see is not evidence-based versus data science, but an appreciation of what data science approaches can provide to the evidence-based for nursing practice.

- Thank you very much. I have a question from Katherine Stansfield. And she asks, "Is open access available for all nurses internationally?"

- Katherine, that's a very good question. Oopen access, any journal, any article that is identified in PubMed Central, or any journal that's identified as an open-access journal in PubMed Central are available internationally with one constraint.

And that is, in countries where there is a constraint on internet use or internet access, those individuals are sometimes not able to get to our resources.

- Thank you. I have one last question for you from Michelle Buck. She says, "Thank you for an excellent and informative presentation. Is the AI process you've described to identify the best resources for data search used by any other data repositories or is it unique to PubMed?"

- That's another good question. We work very closely with our colleagues at Google, who have a very sophisticated search process, and with colleagues around the world who maintain literature and data resources. So the algorithm that we use and the way it is implemented is unique to the NIH. It is documented in an article in Nature from 2018.

The first author is Lu, L-U. And the algorithm is designed to be optimized for an individual search. So it's not defining what is the best article about wound care? Or what is the best article about anxiety management? It says, "Given what you're searching for, what is the best match to your questions?"

This actually relates a little bit to our question earlier about,to what extent does data science compete with evidence-based practice? When we use AI, we're using it for a specific purpose. And we're very cautious not to go beyond that. And that's part of why we do not provide authors, for example, information about how often their particular article turned up in an individual search.

We provided information about the citations. That is once an individual has published a paper and cited an article, we can provide quick access to the number of times an article was cited. But we don't want to see this as a competition, or some kind of a gold star that your article got to pulled more than others. Our approaches to searching and search itself is a wonderful point of research right now, very exciting area of research, where lots of work is being done to try to determine the best most efficient way to search.

My hope in the future is that our search algorithms will become more in the moment being able to answer questions on the fly as they arise.

- Dr. Brennan, thank you so very much for your excellent presentation and the answers to all these really important questions. Thank you again.

- Thank you and good luck with your program today. I appreciate being invited. Thank you.

- Thank you.