Librarian Assistance in the Use of Semantic MEDLINE: Outreach to Scientists

Project Leader
Susan Roy, NLM Associate Fellow

Project Sponsor
Dr. Tom Rindflesch, Lister Hill National Center for Biomedical Communications

Spring/Summer 2012

Table of Contents

Abstract 2

Introduction 3

Methods 6

Results 12

Discussion 20

Conclusions 22

Acknowledgements 23

References 24

Appendix A. 26

Abstract

OBJECTIVE: The purpose of this project was to explore possible roles for librarians in using Semantic MEDLINE (SemMed) as a tool to promote outreach to scientists. Basic scientists are heavy users of library resources but an often underserved population for many libraries. Outreach to scientists by utilizing tools and technology such as SemMed, a program that extracts semantic relations from PubMed citations, can have tremendous potential in fostering collaborations and scientific innovation.

METHODS: Current users of Semantic MEDLINE were interviewed to determine how they use SemMed to uncover and predict literature-based discoveries (LBD). The reciprocity in search behavior that occurred between the user and the computer application was described and analyzed. The discovered principles were then taught to other researchers new to the area of LBD. Scientists from academia, industry and federal institutions were recruited to learn about SemMed by participating in a live, online instructional session. Following the session, connections were maintained with a follow-up survey and persistent outreach, which provided opportunities to gather thoughts about SemMed, the discovery process and collaboration possibilities. Finally, ways SemMed could be utilized as a tool by other librarians to promote science research were hypothesized.

RESULTS: Three current SemMed users were interviewed and the lessons learned were then used to create training tutorials. Twenty-two scientists were recruited and taught how to use SemMed via live, online tutorials. Following the training sessions an online survey was sent out to all participants, and seventeen of the twenty-two scientists (77.27%) responded to the survey. Of those that responded, seven (41.18%) requested SemMed login access information; ten (58.82%) indicated they planned to request access in the near future. All of the scientists indicated that they were able to learn about SemMed from the tutorial. Additionally, scientists indicated they planned to use the tool for LBD research purposes and/or for gaining knowledge about recently published information. In follow-up conversations many scientists were interested in collaborative assistance, and 76% of the survey respondents indicated they would use SemMed for research. Currently, two collaborative projects are underway to assist scientists with SemMed for LBD research.

CONCLUSIONS: This project provided opportunities for the Associate Fellow to learn how to teach basic scientists about SemMed and an opportunity to foster potential future research collaborations. With the methodologies and technologies described here, we show how tools such as SemMed can be used by librarians for making connections to scientists, and we lay the groundwork for the Associate Fellow to continue collaborative efforts with basic scientists and the NLM Semantic MEDLINE research group.

Introduction

The American Library Association® defines "outreach" as any library service that supports access to knowledge and information, especially to meet the needs of an inadequately served target group (1). In academic institutions the basic sciences and research scientists are heavy users of library resources, and yet an often underserved population (2). Few studies have investigated the library resource and service needs of basic scientists (2-6). These studies that have looked at the needs of basic scientists have shown that they will often indicate access to research literature as the most essential library need for their work. While many scientists suggest they do not actively or regularly seek out librarian assistance, two recent studies reported scientists (molecular biologists in these particular studies) needed more tools for mining biomedical literature and large genetic datasets to discover relationships and generate new hypotheses (7-8). Despite having this large library user group with obvious resource needs, librarians still seem to have difficulties in finding ways to provide outreach with basic scientists.

One way that librarians could reach out in support and/or collaboration with scientists is through assistance in literature-based discovery (LBD). LBD is a process that searches for hidden and important connections among information in published literature. Generally there are two types of LBD processes: *open*- or *closed* discovery. *Open* discovery describes LBD used for hypothesis generation, whereas *closed* discovery describes LBD used for hypothesis testing (9).

Swanson, a mathematician and information scientist, was the first to coin the term "literature-based discovery" when he linked two unrelated ideas together based upon common symptoms. In his classic publication, Swanson used published literature to develop the hypothesis that fish oil could be used to alleviate symptoms of Raynaud's Disease (10). At the time it was known that Raynaud's Disease causes blood-platelet aggregation. It was also known that fish oil could thin blood. Swanson then used the published literature to show that fish oil could be used to treat Raynaud's Disease because it lowers blood viscosity. He later went on to publish discoveries linking the role of magnesium deficiency in migraine headaches (11), effects of arginine intake in somatomedin C levels (12), and the protective effects of estrogen in Alzheimer's Disease (13). The first two discoveries were later confirmed experimentally (14).

Because many scientists only have time to focus or specialize in a very narrow area of research, LBD is becoming more relevant as scientific knowledge grows because LBD can help find connections between many different areas of published research (9, 15). Skills needed for LBD include literature and information searching, collecting and disseminating; which are abilities that librarians possess. By combining the librarian skills in literature-based research, and in collaborating with expert scientists, great discovery could occur.

One way that librarians could provide assistance to scientists and promote LBD is with the use of tools such as Semantic MEDLINE (SemMed). SemMed is an advanced information management application for biomedical literature, created and maintained by the Lister Hill National Center for Biomedical Communications (LHNCBC) at the National Library of Medicine® (NLM) at the National Institutes of Health® (NIH) (16-17). SemMed integrates document retrieval, advanced natural language processing (NLP), automatic summarization and visualization into a single Web portal.

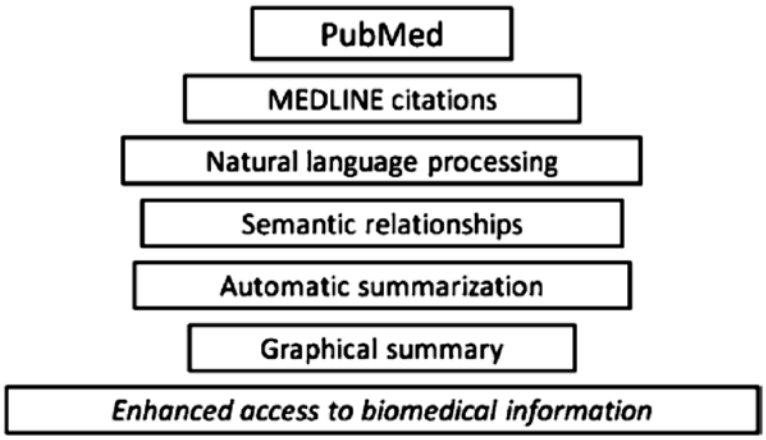


Figure 1. Schematic diagram on Semantic MEDLINE processing (figure from Rindflesch et al., 2011).

Information retrieval in SemMed can be summarized as shown in Figure 1. The application uses PubMed to return MEDLINE citations based on the search query. Semantic relations can be extracted from citations because of (pre-) semantic processing (see Figure 2). And automatic summarization is completed to hone in on the most salient information. Finally, the summarized results are visualized to graphically show links to source text and additional information (see Figure 3). The full version of SemMed is still in development but is freely accessible by requesting login information from Dr. Thomas C. Rindflesch (trindflesch@mail.nih.gov) (16-17).

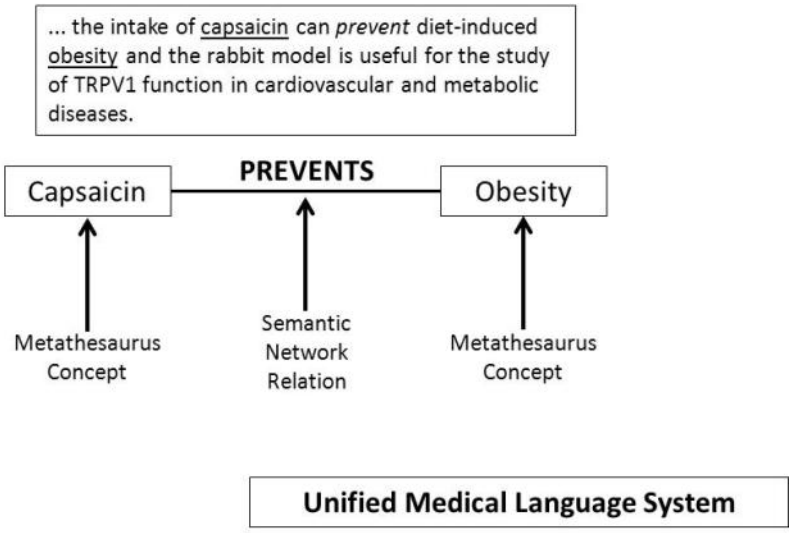


Figure 2. Diagram example of an extracted semantic predication (figure adapted from Rindflesch, et al., 2011).

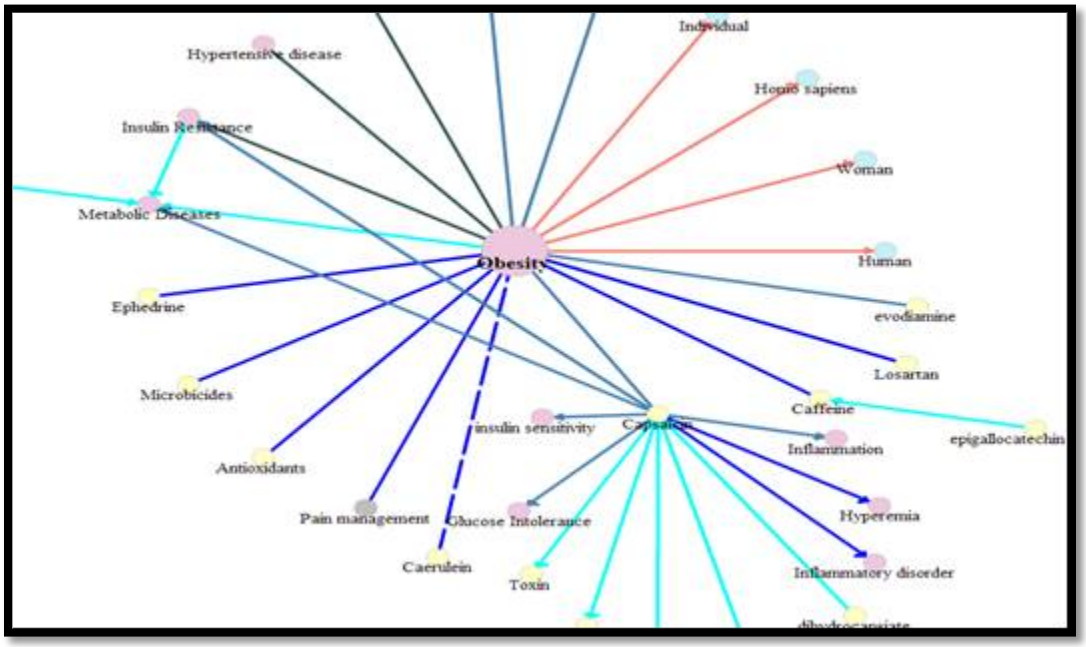


Figure 3. Semantic MEDLINE visualization showing semantic predications extracted from MEDLINE citations retrieved with the search query "capsaicin and obesity".

SemMed can have an impact in biomedical research by exploiting meaning in published literature and assisting scientists in research and scientific discovery. SemMed has already been used for literature-based discovery. Previous studies using SemMed have shown that an altered Per2 clock gene could be linked to some cancers and obesity (17), disrupted sleep quality in aged men could be linked to hypogonadism (18), a hypothesized interaction of norephinephrine, inflammation and chronobiology in depressive disorders (19) and hypothesized functions for the peptide hormone relaxin (16). These are just some examples on how SemMed can be utilized to uncover and predict literature-based discoveries (LBD).

This innovative application is still being developed, but it is ready for librarians and scientists to use for LBD. However, because of the complexity and the relative novelty of semantic technologies, additional assistance in understanding the system may be needed. Additionally, librarians are in constant need to learn new approaches to provide outreach and services to basic and clinical scientists. Therefore, the overall purpose of this project was multi-faceted. By teaching scientists about SemMed and how to use the tool, the Associate Fellow could learn how to teach basic scientists and have an opportunity to foster potential future research collaborations. Also the methods and results discovered in this project could be used to explore the librarians' role in using Semantic MEDLINE as a tool to promote outreach to scientists. Finally, these efforts could assist in the overall objective to increase awareness of Semantic MEDLINE to librarians and biomedical scientists.

Methods

To accomplish the many goals of this project, the Associate Fellow first learned about SemMed, then created an online tutorial and provided outreach to scientists. It was hypothesized that these results would then provide evidence for a librarian's role in using Semantic MEDLINE as a tool to promote outreach to scientists.

Semantic MEDLINE Observational Studies

Prior to teaching scientists how to use SemMed for literature discovery, it was necessary to first learn how the tool is developed, maintained and used. For the first month of the Spring Project phase the Associate Fellow read papers, discussed with colleagues, and joined the SemMed research group. To learn how to use SemMed, the Associate Fellow shadowed three NLP scientists from the SemMed group who have used the tool for literature-based discovery. These three SemMed users were observed for 1-2 hours each, as they completed search tasks on SemMed. Particular attention was paid to how they used SemMed to develop research ideas, search for information and answer questions. The knowledge gained from observing the user behaviors was incorporated into creating a SemMed tutorial/instructional that would be most effective for teaching scientists.

Semantic MEDLINE Outreach and Instruction to Scientists

A presentation was developed that could be used to quickly and easily teach scientists. The tutorial was developed and tested on one pilot participant, resulting in minor modifications, including a comparison of SemMed to PubMed, as well as definitions of terms frequently used in semantic research. The final version of the PowerPoint presentation consisted of five slides that provided detail and background context about SemMed but without the difficult semantic technology jargon. The Associate Fellow made a priority to keep the SemMed tutorial simple and easy to understand. All sessions utilized the same PowerPoint presentation and SemMed search demonstration.

The tutorial began by introducing SemMed as a Web-based application that integrates PubMed searching, advanced natural language processing, automatic summarization, and visualization in a single portal (see Figures 4). Scientists were shown a typical PubMed citation and then shown how SemMed will text mine the title and abstract of citations (see Figure 5). The participants were given more information about how the system utilizes advanced natural language processing to create semantic relationships in RDF triples. A triple was explained to consist of a subject (the subject or what the object is describing), predicate (the relationship) and an object. Therefore, SemMed was described as a system that essentially extracts semantic predications, or statements on how one object relates to another, from PubMed citations (see Figure 6).

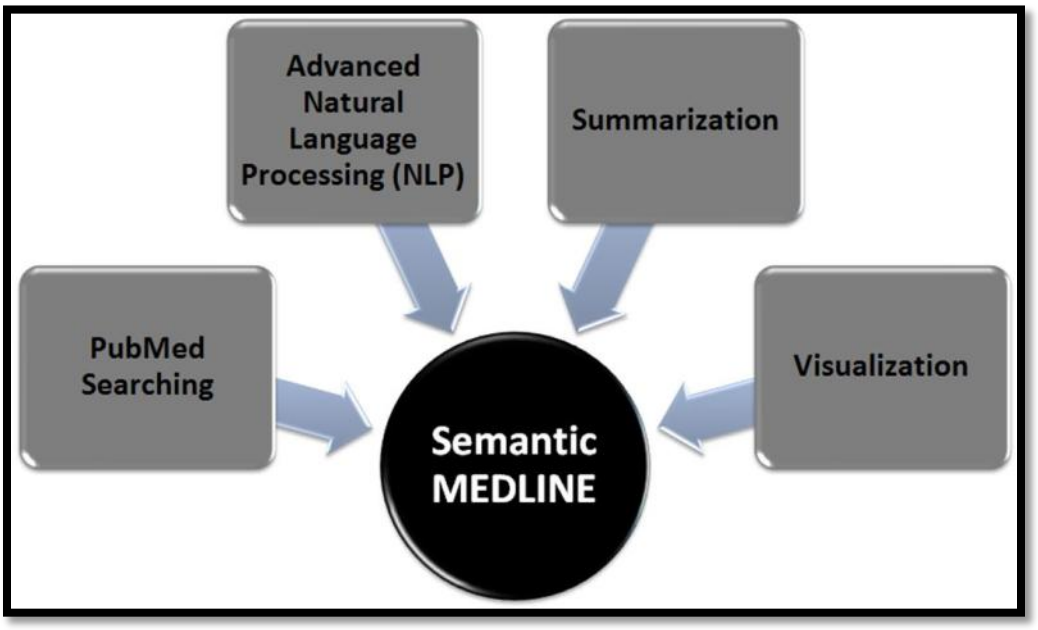


Figure 4. Screenshot of the introductory slide of the tutorial. The purpose of the introduction slide was to illustrate the integrated functions (i.e., PubMed searching, natural language processing, summarization and visualization) of SemMed.

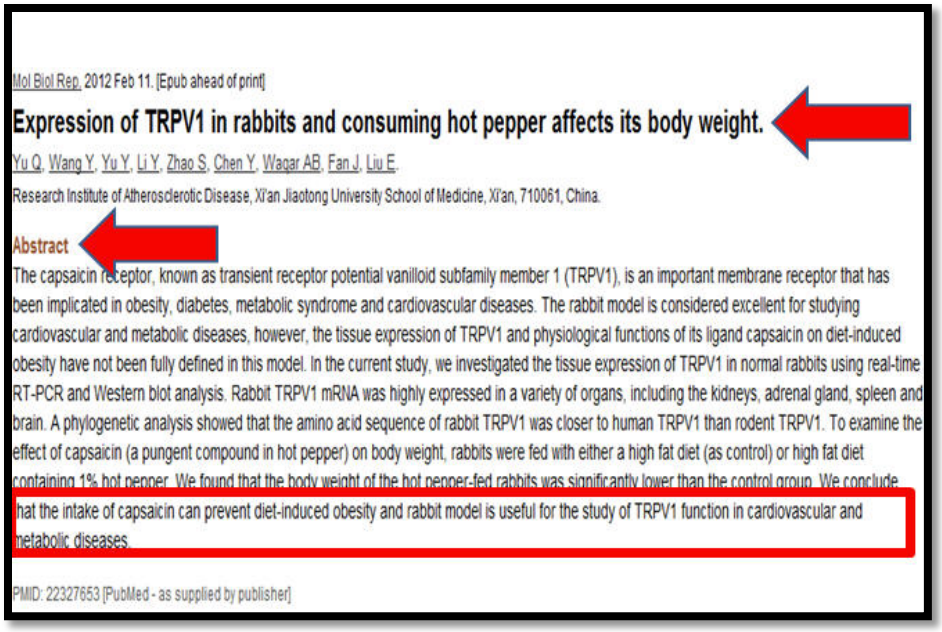


Figure 5. Screenshot of the tutorial slide showing a text mined citation and illustrating how semantic relations are extracted in SemMed.

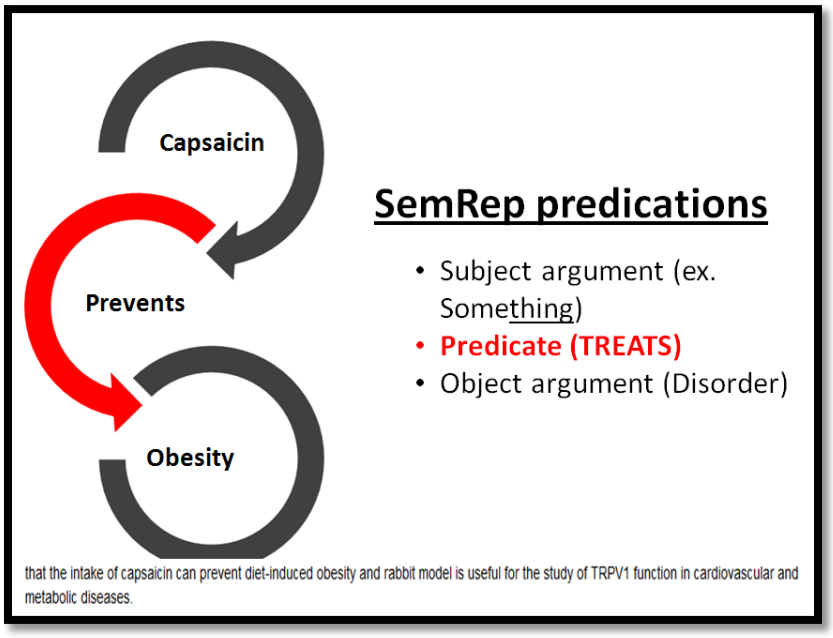


Figure 6. Screenshot of the tutorial slide illustrating a semantic predication (subject-predicate-object), or relation statement extracted from the PubMed citation.

Following the brief introduction, participants were given a demonstration of SemMed. During the demonstration the same search (ex. "Parkinson* AND resveratrol") was conducted in both SemMed and PubMed because the scientists participating in the tutorials were assumed to be heavy PubMed users. By showing the same search in both systems, it was hypothesized that it would provide evidence on how SemMed is both similar and different from PubMed. That is, by first showing how the citations results were similar in PubMed and SemMed (60 citations in our example query), we were then able to show how SemMed extracts not only citations but also predications from the citations (434 predications from the 60 citations in our example) (see Figures 7 and 8).

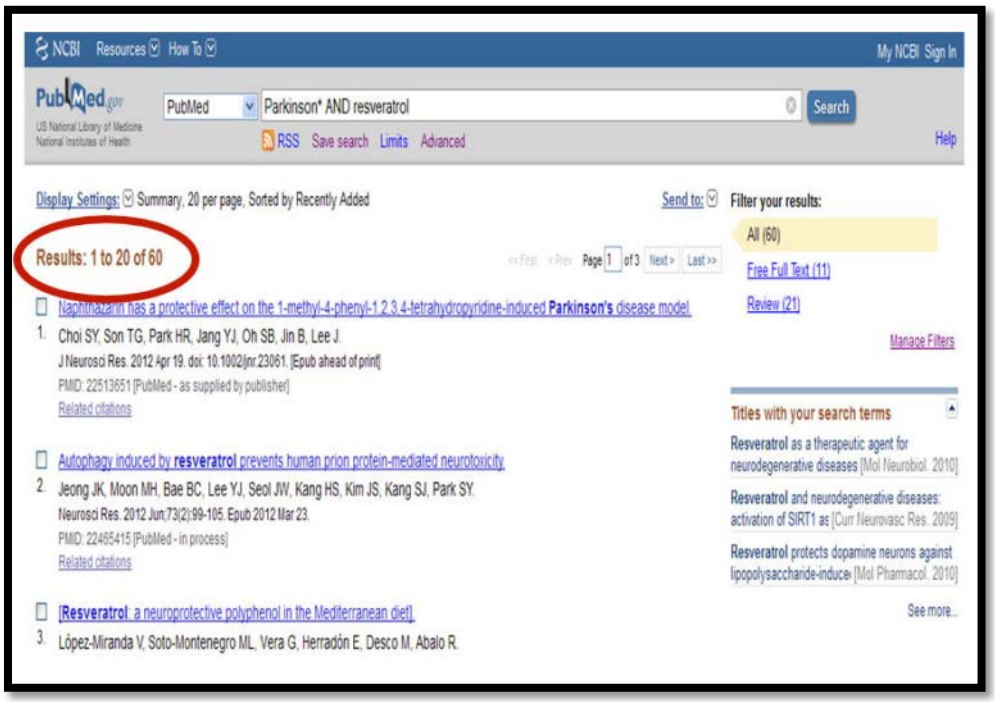
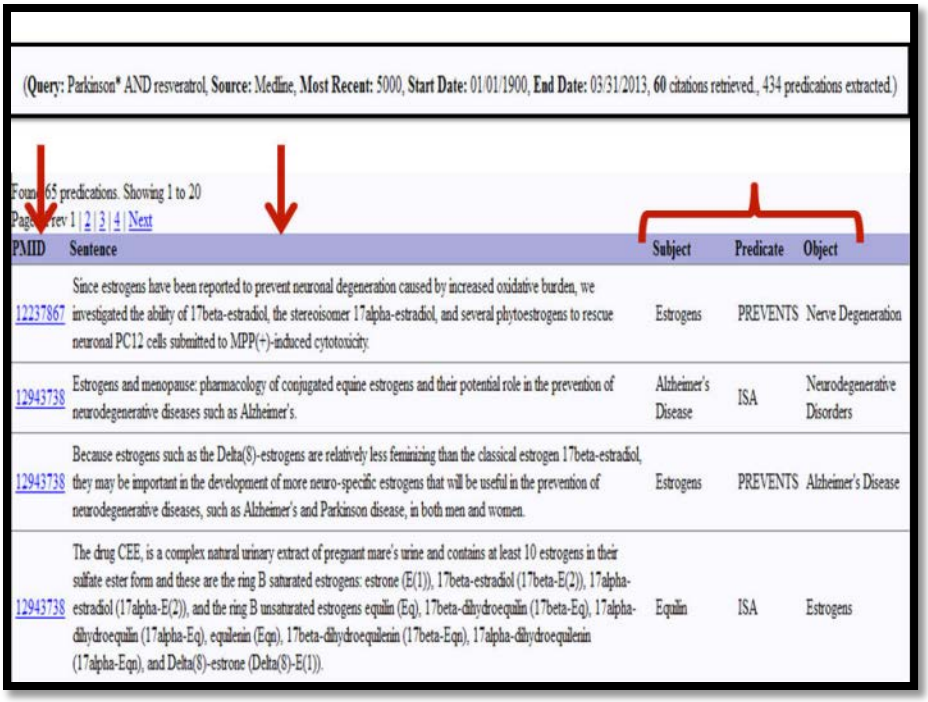


Figure 7. Screenshot of the tutorial slide illustrating a simple search query (Parkinson* AND resveratrol) in PubMed. 60 citations were retrieved.



(**Query:** Parkinson* AND resveratrol, **Source:** Medline, **Most Recent:** 5000, **Start Date:** 01/01/1900, **End Date:** 03/31/2013, **60** citations retrieved., 434 predications extracted.)

Found 65 predications. Showing 1 to 20
Page: Prev 1 | 2 | 3 | 4 | Next

PMID	Sentence	Subject	Predicate	Object
12237867	Since estrogens have been reported to prevent neuronal degeneration caused by increased oxidative burden, we investigated the ability of 17beta-estradiol, the stereoisomer 17alpha-estradiol, and several phytoestrogens to rescue neuronal PC12 cells submitted to MPP(+)-induced cytotoxicity.	Estrogens	PREVENTS	Nerve Degeneration
12943738	Estrogens and menopause: pharmacology of conjugated equine estrogens and their potential role in the prevention of neurodegenerative diseases such as Alzheimer's.	Alzheimer's Disease	ISA	Neurodegenerative Disorders
12943738	Because estrogens such as the Delta(8)-estrogens are relatively less feminizing than the classical estrogen 17beta-estradiol, they may be important in the development of more neuro-specific estrogens that will be useful in the prevention of neurodegenerative diseases, such as Alzheimer's and Parkinson disease, in both men and women.	Estrogens	PREVENTS	Alzheimer's Disease
12943738	The drug CEE, is a complex natural urinary extract of pregnant mare's urine and contains at least 10 estrogens in their sulfate ester form and these are the ring B saturated estrogens: estrone (E(1)), 17beta-estradiol (17beta-E(2)), 17alpha-estradiol (17alpha-E(2)), and the ring B unsaturated estrogens equilin (Eq), 17beta-dihydroequilin (17beta-Eq), 17alpha-dihydroequilin (17alpha-Eq), equilenin (Eqn), 17beta-dihydroequilenin (17beta-Eqn), 17alpha-dihydroequilenin (17alpha-Eqn), and Delta(8)-estrone (Delta(8)-E(1)).	Equilin	ISA	Estrogens

Figure 8. Screenshot of the tutorial slide illustrating the same simple search query (Parkinson* AND resveratrol) in SemMed. Shows the same 60 citations retrieved in PubMed, and the 434 predications extracted in SemMed.

The visualization function of SemMed was then demonstrated and it was explained how SemMed could bridge the gap between text and meaning by extracting the semantic predications from the retrieved citations and displaying the connections between topics graphically (see Figure 9). Following the tutorial and demonstration participants were given the opportunity to try a search in SemMed, ask questions, and converse freely with each other.

To recruit scientists to participate in the project, the Associate Fellow used Facebook, Twitter and Google+ (G+) to connect with personal scientist contacts. Social networks were then utilized by requesting friends to spread the word through their own contacts for additional secondary and tertiary (to the Associate Fellow) research scientist contacts. Participation was voluntary and no compensation was offered (other than access to the application or additional assistance with SemMed use).

Four Adobe Connect sessions were conducted during the months of May – June, 2012. One session via Skype™ and an additional one session via G+ Hangout (Web-based video conferencing programs) were conducted in July 2012.

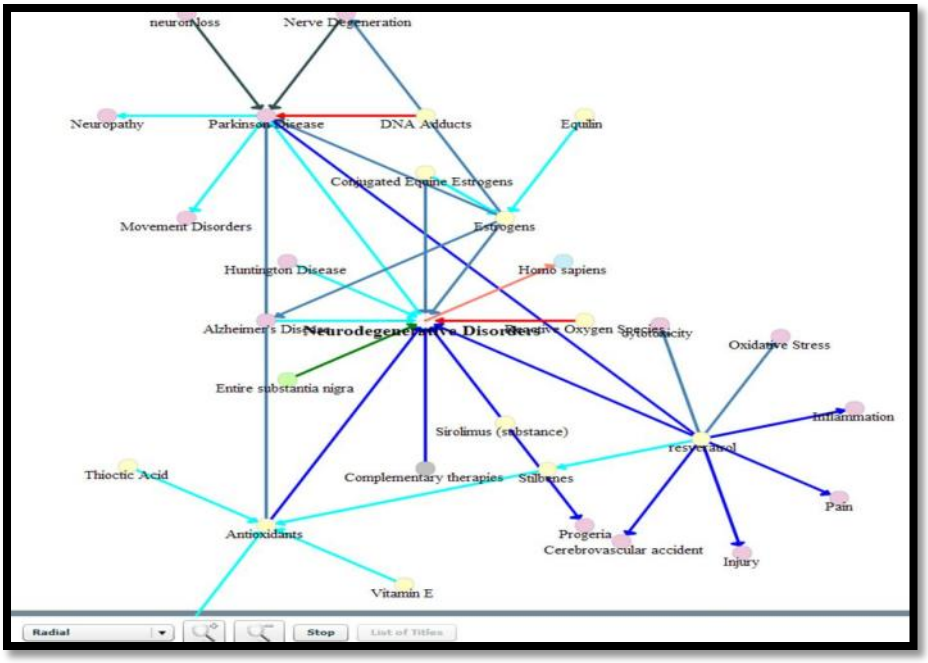


Figure 9. Screenshot of the tutorial slide illustrating SemMed visualization of predications extracted from the search query (Parkinson* AND resveratrol).

Follow-up with Scientists

To investigate the effectiveness of the tutorial and to follow-up for potential collaborative projects, a survey was sent out to the scientists 1-2.5 weeks after SemMed tutorial participation. The questionnaire was developed by the Associate Fellow and then revised following peer-review by three colleagues. The survey was administered electronically using Survey Monkey® (http://www.surveymonkey.com/), a free online-based survey tool, by sending the link via email. The survey contained nine open response and/or Likert scale rated questions (see Appendix A). The survey was left open and responses were collected for an additional month after the last tutorial session was completed. Exempt status was sought and granted from the NLM/NIH Institutional Review Board and granted for the survey.

Additional follow-up correspondence with the scientists was conducted informally using email, Facebook, Twitter, phone calls and Skype.

Results

Semantic MEDLINE Observational Studies

To understand the underlying principles of Semantic MEDLINE, the Associate Fellow joined the SemMed research group and found this unique perspective of being a member of the team invaluable in learning about the tool. Additionally, observation sessions provided opportunity to investigate the reciprocity in search behavior between three proficient SemMed users and the computer application. Table 1 provides a summary from the observation sessions.

During the observation sessions it was noted that all three NLP scientists utilized PubMed and SemMed interchangeably during their topic exploration. But how the three scientists used the two applications differed. One used SemMed more for exploring and for seeking out the less known connected concepts. PubMed was then used to read the citations from those connections. Another NLP scientist used SemMed to browse the more highly connected nodes, and then used PubMed to verify citations and the quality of the research. These different ways of exploring information show the versatility of the system.

It was interesting to learn (during the observation sessions) that all three NLP scientists utilized a mind mapping software program (i.e., MindManager™ and XMind) to organize information as it was learned. When questioned, the NLP scientists explained that these tools provide assistance in organizing information and for brainstorming. This idea was found to be noteworthy because the proficient SemMed user values the visualization aspect of SemMed and yet also depends on additional programs to help visualize and organize the information as it is learned.

It was an advantage to study the three NLP scientists' discovery process because all three users searched SemMed in a different way. Two of the observed users demonstrated SemMed with an *open* discovery process and one demonstrated a *closed* discovery process. Because each user had their own way of searching and exploring SemMed, it was intriguing to see how SemMed helped to broaden their knowledge on a topic by pure exploration of the summarized semantic relations. All three users demonstrated that they use the tool to gain knowledge (or additional knowledge) about questions and ideas, regardless of the *open*- or *closed*-discovery type of search. For example, all three users were open to the ideas and concepts that were new to them by actively seeking the connections that were surprising. This idea that SemMed can be used to discover novel and surprising concepts in the literature is a critical application for the program. The observation of three users, and their interaction with the system, attests to the ingenious qualities of the SemMed program and shows how even experts within a particular research field domain can learn new information by using SemMed. This knowledge from observing the three proficient SemMed users was then used to create an online tutorial for scientists.

SemMed User 1	SemMed User 2	SemMed User 3
Uses SemMed for literature/idea exploration but would switch between PubMed and SemMed throughout the interview.	Prior to search, carefully articulated a question that was to be answered, then formulated a search query, and then would systematically search SemMed.	During the search the user would seek out the concepts that were most highly connected and then read those relations.
During the interview the user systematically clicked and read all relation links and concepts in a graph.	The user would seek out what was already known (to the user) and then would build on those assumptions.	Utilized additional resources (i.e., MindManager and XMind) to assist in the brainstorming and mind mapping of the newly learned ideas and concepts.
Would start to build own knowledge-base from the summarization and visualization and tried to understand all the connections of the nodes.	Constantly revised the question and created new, focused search queries to answer the question.	Used SemMed more to hone in on an idea that was in question as if trying to find answers to a hypothesis.
If the user did not know a concept that was displayed, the user would use Google and PubMed to learn more.	Paid particular attention to the quality of the papers from which the predications were extracted. Used PubMed and SemMed almost interchangeably.	

Table 1. Summary of three proficient SemMed user interviews.

Semantic MEDLINE Outreach and Instruction to Scientists

22 scientists were recruited to participate in the Semantic MEDLINE tutorials. Of these 22 scientists, two were from industry, one was from federal and 19 were from academic settings. Of the 19 academic scientists, there were seven graduate students, one technician, six post-doctoral students, and five professors (see Figure 10). 14 of the 22 participants had PhD degrees, five had Master Degrees, and of those without PhD degrees all but one were currently enrolled in a graduate program leading to a PhD.

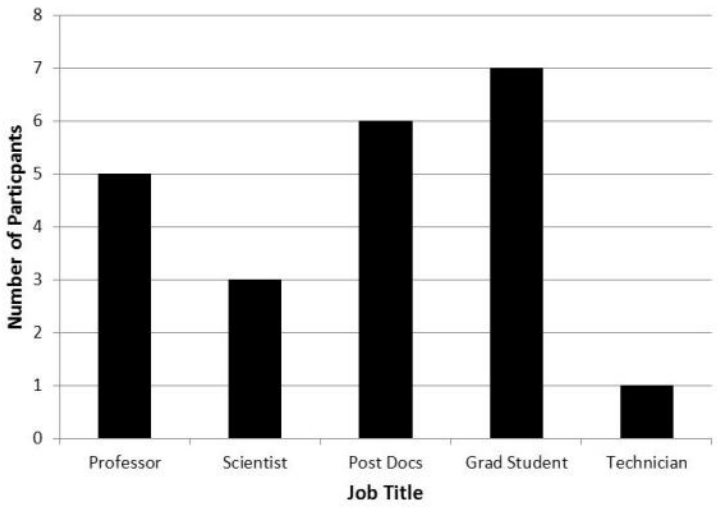


Figure 10. Job demographics for SemMed tutorial participants.

A total of six online sessions were conducted, four Adobe Connect sessions (20 participants), one Skype (one participant) and one G+ Hangout (one participant). Table 2 shows the distribution of the participants during the various SemMed tutorial sessions.

SemMed Instructional Session	Number of Participants
Adobe Session 1	7
Adobe Session 2	6
Adobe Session 3	3
Adobe Session 4	4
Skype	1
G+ Hangout	1

Table 2. Number of participants during each Semantic MEDLINE tutorial session.

The online tutorials took an average of 35 minutes. Approximately 7 minutes were spent on the SemMed introduction and description on how the application extracts semantic predications from PubMed citations. The next 15 minutes were a demonstration of a search in PubMed and SemMed. And the last 10-15 minutes of the session were dedicated to assisting the group or individuals with their own searches, answering questions and fostering discussion amongst the participants.

Qualitative evaluation of the tutorial sessions favored the positive side with nearly all scientists interested in using SemMed for research purposes. A few of the scientists also indicated that the tool would be helpful for post-doctoral students and for new/young researchers/investigators. A number of the participants alluded to the fact that tools, such as SemMed, could assist with developing innovative hypotheses, and others suggested it would assist in searching through the literature for new “undiscovered” ideas. Additional examples of comments from participants can be found in Table 3.

Tutorial Comments
“Semantic MELDINE offers great search capability and the walk-through was the most helpful part of the tutorial.”
“I like how short but informative this tutorial was – thank you!”
Research Comments
“Nice streamlined process of searching.”
“This would be helpful for when I’m writing my book chapter. I will need to come up with ideas and hypotheses and this will really be helpful.”
“The search you have demonstrated was interesting; I would never think to look at some of these genes with that disease.”
“I can see how this tool will help me organize the vast number of articles that I’m supposed to read.”
Negative Semantic MEDLINE Comments
“I can see the point, but this is not helpful for certain types of scientists.”
“I wish it would search methods – that would be really helpful.”
“I’m still confused about the UMLS concept summarization thing.”
Positive Semantic MEDLINE Comments
“Cool!”
“Thank you (to NLM) for developing this tool.”
“Great tool.”
“I think this will really help science.”
“Can I get a login today?”
“I heard about this tool from a colleague who took your last tutorial and I was interested in learning more – this is great! Thank you!”
“I wish I knew about this system before I wrote and defended my thesis.”
“This would have made coming up with ideas for graduate school easier.”
“I love how NLM is helping scientists – this is so cool!”

Table 3. Example of comments made by scientists during the tutorial sessions.

Follow-up with Scientists

One- to two and a half weeks after the tutorial sessions, a survey was sent via email to the scientists to investigate the effectiveness of the online SemMed tutorial (see Appendix A for survey). 17 of the 22 Semantic MEDLINE tutorial participants completed the survey for a response rate of 77.27%. All 17 (100%) survey respondents indicated that they were able to learn about SemMed from the online tutorial.

Following the online tutorials, the scientists were told that if they would like (free) login information for SemMed they needed to send an email request to the Associate Fellow. Of the 17 survey respondents, seven (59%) had already requested for SemMed login and the other 10 (41%) respondents indicated they had not yet requested login but they plan to in the future. None of the survey respondents said they did not plan to request login information.

For those respondents that had requested SemMed login information, they were asked to indicate how often they had used SemMed since the tutorial. Out of the 7 survey respondents that had requested login information, one participant had not yet used SemMed (14%), four had used SemMed 1-2 times (57%), two had used SemMed 3-5 times (29%) but none had used SemMed six or more times (see Figure 11).

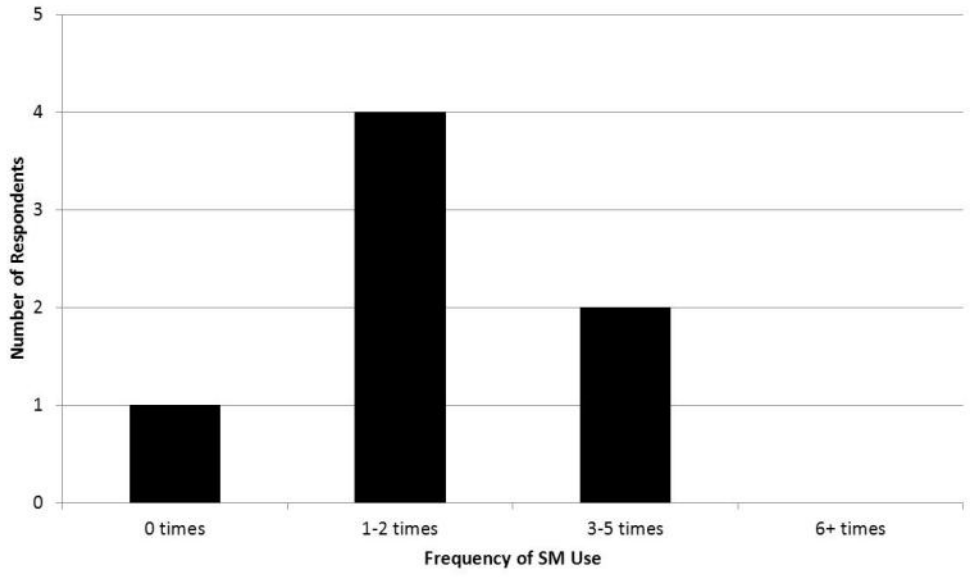


Figure 11. Displays the number of times survey respondents indicated they have used SemMed since the tutorial and receiving login information.

Next we wanted to investigate how likely scientists might use SemMed for research. We asked them to rate how likely (*Very Likely, Somewhat Likely, Unsure, Somewhat Unlikely, Very Unlikely*) they themselves would use the tool for research, finding or facilitating collaborations, studying/gaining knowledge, biomedical literature exploration or hypothesis generation (see Figure 12). Overall there was positive feedback; a majority of the scientists indicated it was *Very Likely* or *Somewhat Likely* that they would use SemMed for many aspects of research. Scientists indicated they would be *Very Likely* to use SemMed for research (76%), studying or gaining knowledge (82.35%), or for biomedical literature exploration (58.82%). In fact, SemMed for research, studying/gaining knowledge and for biomedical literature exploration was rated with no less than a very positive *Very Likely* or *Somewhat Likely* rating. Very few scientists rated on the neutral or negative side of the continuum with *Unsure* (finding or facilitating collaborations, 35.29% and hypothesis generation, 17.65%) and *Somewhat Unlikely* or *Very Unlikely* (hypothesis generation, 17.65% for both). Overall these results show there was mostly positive affirmation for SemMed use by scientists for a number of different research purposes.

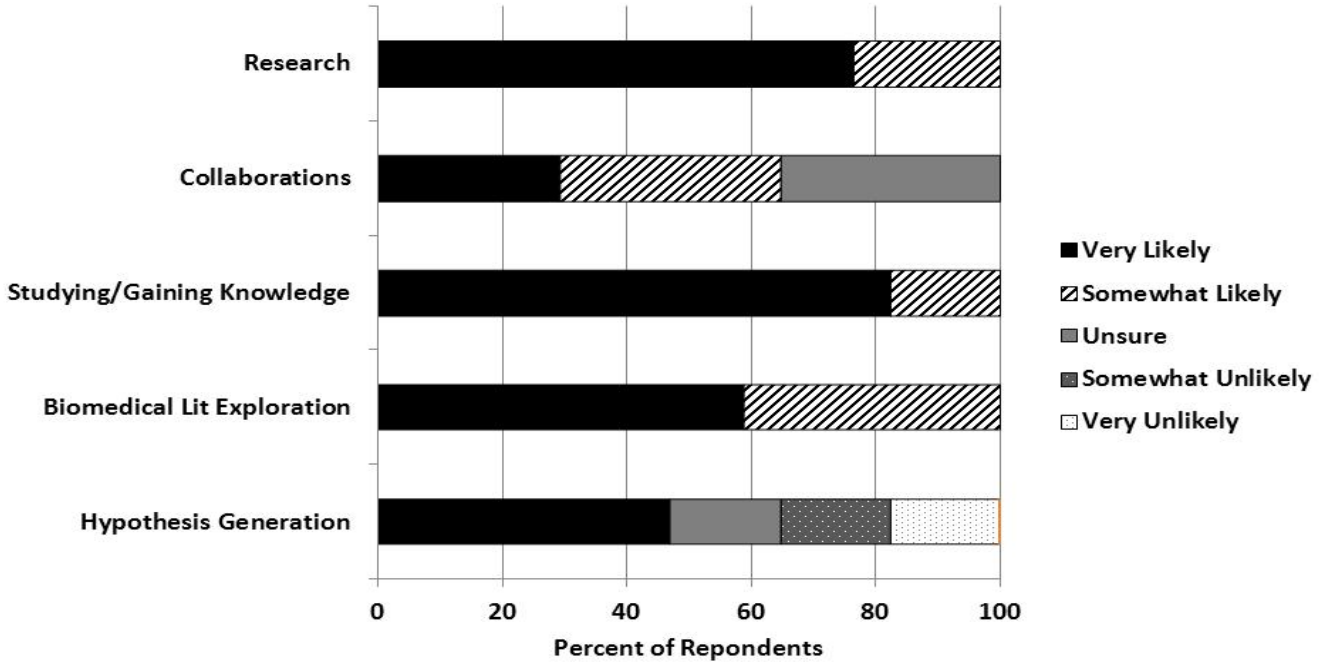


Figure 12. Displays the percent of responses for "How likely would you use SemMed for - Research, Finding/Facilitating Collaborations, Studying/Gaining Knowledge, Biomedical Literature Exploration and Hypothesis Generation".

Scientists were also asked to indicate whether they thought SemMed would be a good tool for others in research (see Figure 13). All 17 respondents indicated SemMed would be a good tool for other academic scientists, graduate students, physicians and librarians (100%). Most respondents also indicated that SemMed would be useful for industry scientists, medical students and other health care professionals (82.35%). Undergraduate students were rated the least to find SemMed useful (*Yes* 52.94% and *Maybe* 47.06%). None of the scientists indicated *No* in any category, suggesting SemMed could be useful to most people interested in biomedical research and knowledge.

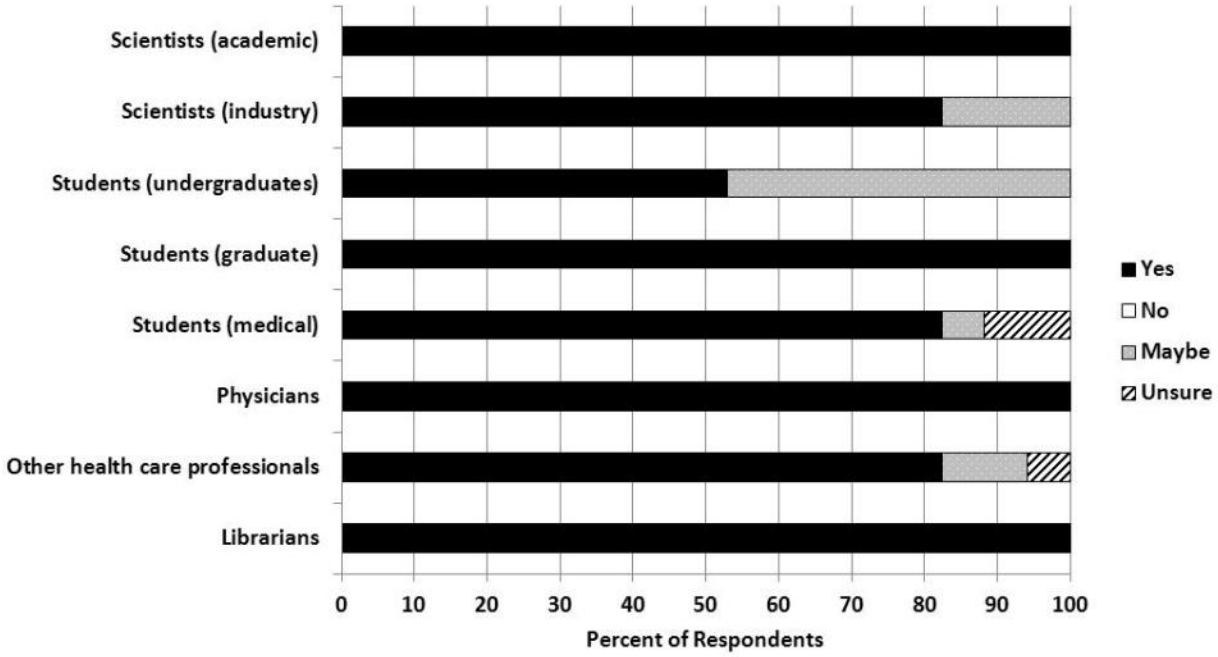


Figure 13. Survey respondents were asked to indicate *Yes*, *No*, *Maybe* or *Unsure* as to whether they thought SemMed would be a useful resource for Scientists (academic or industry), Students (undergraduates, graduates or medical), Physicians, Other Health care professionals or for Librarians.

The final question in the follow-up survey asked scientists whether they thought they might need additional assistance with the use of SemMed (see Figure 14). Most (53%) scientists indicated *No*, they did not need any additional assistance. Another 29% said *Maybe*, 12% said they were *Unsure*, and 1% said *Yes*, they would like additional assistance with SemMed. After completion of the online tutorial sessions, additional follow-up with the scientists via phone calls, emails, Twitter and/or Skype conversations was completed. All contacted scientists welcomed the opportunity to talk about their current research and strategizing how SemMed and PubMed searches could assist in furthering discovery. Additionally, since the tutorials one scientist has

contacted the Associate Fellow regularly for assistance on query strategizing and another two scientists are in the process of writing (different) book chapters. Both have indicated they would like assistance with using SemMed to generate new research topics and ideas for their book chapters. Another scientist was writing a grant and wanted assistance in exploring the literature using SemMed for hypothesis and specific aim ideas. Overall these follow-up efforts have proven beneficial and refute the suggestion that scientists do not want additional assistance. These results show that while many scientists do not regularly seek out librarian assistance, they are not opposed to the service when approached.

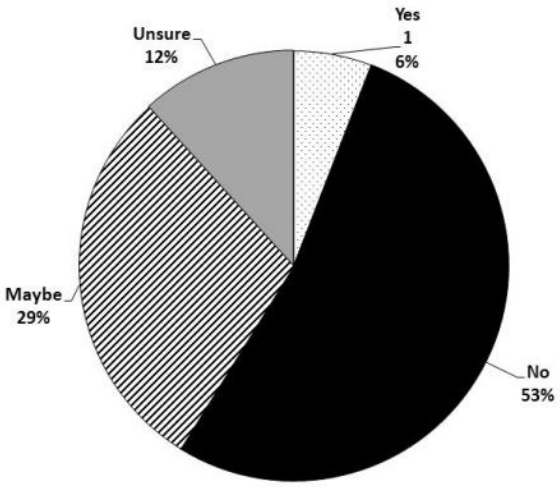


Figure 14. Scientist response to the question: Will you need additional assistance with the use of Semantic MEDLINE?

Discussion

Here we report on a project where tutorials were created and conducted to show how librarians could use SemMed for outreach to scientists. We aimed to teach scientists about SemMed and analyze their thoughts about using the tool for research. The project results suggest that the scientists who participated in the tutorial were able to easily learn about SemMed and were interested in using the tool for research. Many of the participants requested, or indicated they planned to request, login information to use the application for research. Overall these results show that teaching scientists about SemMed is a feasible undertaking, and these results are promising for librarians in providing outreach to scientists.

Scientists, especially those in the basic and integrated fields, are an important target group because they tend to be heavy users of library resources (2-3, 5-6). Many libraries offer in-

context informationists or subject-specific library liaisons to serve the basic sciences and yet there are still limited connections to basic research groups. One possible reason for the disconnect could be because many scientists tend to believe they are self-sufficient in finding their own information resources (3). Our follow-up survey results show similar findings to what others have reported in regards to basic and academic scientists information assistance needs. 53% of our survey respondents indicated they did not think they would need additional assistance with SemMed searching. However we found that when contacted, our scientist participants were willing to converse about current research and possible SemMed use. The Associate Fellow was blind to the respondents of the survey, so it is unknown who indicated they would not need additional librarian assistance with SemMed. The survey respondents who indicated they did not need additional assistance might or might not be the same scientists contacted in follow-up communication. However for those scientists that have been contacted post-tutorial, it is apparent that additional assistance has been necessary, helpful, or appreciated.

The results also reveal scientists' interest in utilizing SemMed as a tool for research and discovery. All scientists surveyed indicated they would likely use SemMed in some way for their own research. This enthusiasm is reflective of recent work where biologists were identified as a group in need of additional tools for mining literature for discovery (7). Many areas of scholarly research have been rapidly changing due to the advances in technology where the boundaries between the sciences have become increasingly blurred and are more interdisciplinary, integrated and comparative. And the increased numbers of available published articles and journals in biomedicine have only added to this rapid change in research. Academic science research is one area that has been greatly affected by these changes, and the rapid advances in technology and information can elude even the most esteemed scientist. However, after teaching scientists about the biomedical application SemMed, it was apparent that the scientist could easily use the resource. The novelty of this system in providing access to not just biomedical citations, but also meaning through visualized semantic relationships was impressive to many of the scientists. Many of the scientists' commented on the usability of SemMed as a tool to assist in keeping up-to-date on published research and trends. And tutorial participants were surprised how easy the Web-based program was to use for searching and extracting biomedical facts. These results support the idea of librarians and scientists using SemMed for exploiting discovery.

Another promising result from this project was the potential for SemMed outreach to foster collaborations. During follow-up conversations with the scientists, many were open to continued collaborative efforts and receiving assistance with SemMed. These results were not surprising because others have reported that basic scientists are very collegial, but they tend to primarily interact with coworkers in their own laboratories and research colleagues at other institutions (2). Scientists value interdisciplinary research and see it as important for the advancement of knowledge and science (20). For the librarian, these data show the potential in collaborating in integrated research efforts. In fact, studies have reported that librarians found most success and communication with research groups when the librarian was recognized as an "everyday presence" or as part of the research environment (21). The idea of using SemMed to foster collaborations and integrate oneself as the information specialist, or librarian for the research

team, could be crucial in providing outreach to scientists. Here we show that teaching scientists about SemMed has provided avenues to potentially becoming a member of the research group.

Finally, it is important to stress the effectiveness of short and simple tutorials. Previous studies have shown that scientists want short instructional sessions that are offered multiple times and online (4). Studies have also mentioned the difficulties in providing library services to researchers located distantly from the library (3, 6). Here we successfully utilized multiple modes of online tutorial sessions (i.e., Adobe Connect, Skype or G+) and offered these sessions multiple times during a two-month span. We found that offering these tutorial sessions online increased the number of participants at multiple institutions, which in turn, provided opportunities for the scientists to meet virtually and talk about their research to one another.

Conclusions

Here we show how SemMed can be used effectively in providing outreach to basic scientists and how librarians can use the tool to promote LBD collaborative research. Our recommendations to librarians interested in using SemMed or other technology to provide outreach and research collaboration with scientists should consider:

- Identifying the right tools that can utilized by the researcher.
- Identifying the target research group.
- Creating short, simple tutorials and providing ample opportunities for scientists to participate.
- Communicating often with scientists.
- Being open to new learning and research opportunities.

Overall this project provided excellent opportunities for the Associate Fellow to learn how to teach basic scientists about SemMed and foster potential future research collaborations. With the methodologies and technologies described here, we show how tools such as SemMed can be used by librarians for making connections to scientists with outreach and LBD research potential. Additionally, this project laid the groundwork for continued collaborative efforts with the LHNCBC Semantic MEDLINE research group at NLM.

Acknowledgements

I would like to thank my project sponsor, Dr. Thomas Rindflesh, for being willing to work with me to create an Associate Fellow Spring Project. Also I am greatly indebted to Dr. Rindflesch for all of his assistance and advice throughout the research process and I look forward to continuing our collaborative endeavors!

A big thank you to Dr. Kathel Dunn, Dr. Rindflesch and Jessi Van Der Volgen for reviewing the survey and providing details to ponder in my outreach and instruction to scientists. And thanks to Dr. Dunn and Dr. Rindflesch for gracious edits and reviews of reports drafts, conference abstracts and eventual manuscript.

I am also grateful to Dr. Donald A.B. Lindberg, Betsy Humphreys, and the National Library of Medicine for continued support of the Associate Fellowship Program. And I would like to thank Sheldon Kotzin, Joyce Backus for support throughout the fellowship year.

And I would like to thank Dr. Dunn for her guidance and support throughout the year. To my fellow 2011-2012 Associate Fellows, Bethany Harris, Jessi Van Der Volgen and Michele Mason-Coles for comradeship throughout the Fellowship year. And last but not least, a big thank you to my husband, Ryan Spaulding.

References

1. ALA: Office for Literacy and Outreach Services [Internet]. American Library Association; c2012 [cited 2012, August 19]. Available from: http://www.ala.org/offices/olos/.
2. Haines LL, Light J, OMalley D, Delwiche FA. Information-seeking behavior of basic science researchers: implications for library services. J Med Libr Assoc. 2010;98(1):73-81.
3. Grefsheim SF, Rankin JA. Information needs and information seeking in a biomedical research setting: a study of scientists and science administrators. J Med Libr Assoc. 2007;95(4):426-434.
4. Kuruppu PU, Gruber AM. Understanding the information needs of academic scholars in agricultural and biological sciences. JAL. 2006;32(6):609-623.
5. Hemminger BM, Lu D, Vaughan KTL, Adams SJ. Information seeking behavior of academic scientists. ASIS&T. 2007;58(14):2205-2225.
6. Tennant MR, Cataldo TT. Development and assessment of specialized liaison librarian services. Med Ref Serv Q. 2002; 21(2):21-37.
7. Jensen LJ, Saric J, Bork P. Literature mining for the biologist: from information retrieval to biological discovery. Nat Rev Genet. 2006;7(2):119-129.
8. MacMullen J, Denn O. Information problems in molecular biology. J Am Soc Inf Sci. 2005;56(5):447– 456.
9. Weeber M, Bruza P. Literature-based Discovery. Information Science and Knowledge Management series. Berlin: Springer; c2008. Chapter 1, Literature-based discovery? The very idea; p.3-12.
10. Swanson DR. Fish oil, Raynaud's syndrome, and undiscovered public knowledge. Lib. Q. 1986;56:103–118.
11. Swanson DR. Migraine and magnesium: eleven neglected connections. Perspect Biol Med. 1988;31(4):526–557.
12. Swanson DR. Medical literature as a potential source of new knowledge. Bull Med Libr Assoc. 1990;78(1):29-37.
13. Smalheiser NR, Swanson DR, Linking estrogen to Alzheimer's disease: An informatics approach. Neurology. 1996;47:809-810.
14. Smalheiser NR, Swanson DR. Assessing a gap in the biomedical literature: magnesium deficiency and neurologic disease. Neurosci Res Commun. 1994;15:1-9.
15. Kostoff RN. Literature-related discovery (LRD): Introduction and background. Technol Forecast Soc Change. 2008;75:165–185.
16. Kilicoglu H, Fiszman M, Rodriguez A, Shin D, Ripple A, Rindflesch TC. Semantic MEDLINE: A web application for managing the results of PubMed searches. Proceedings of the Third International Symposium for Semantic Mining in Biomedicine:2008:69–76.
17. Rindflesch TC, Kilicoglu H, Fiszman M, Rosemblat G, Shin D. Semantic MEDLINE: An advanced information management application for biomedicine. Inf Serv Use. 2008;31:15-21.
18. Miller CM, Rindflesch TC, Fiszman M, Hristovski D, Shin D, Rosemblat G, Zhang H, Strohl KP. A closed literature-based discovery technique finds a mechanistic link

between hypogonadism and diminished sleep quality in aging men. Sleep. 2011;35(2):279-285.
19. Wilkowski B, Fiszman M, Miller CM, Hristovski D, Arabandi S, Rosemblat G, Rindflesch TC. Graph-based methods for discovery browsing with semantic predications. AMIA Annu Symp Proc. 2011;1514-1523.
20. Nair KM, Dolovich L, Brazil K, Raina P. It's all about relationships: a qualitative study of health researchers' perspectives of conducting interdisciplinary health research. BMC Health Serv Res. 2008;8:110–20.
21. Haglund L, Olsson P. The impact on university libraries of changes in information behavior among academic researchers: a multiple case study. J Acad Libr. 2008;34(1):52–9.

Appendix A.

Semantic MEDLINE Follow-up Survey for Scientists

1. Were you able to learn about Semantic MEDLINE from the Adobe Connect session?

Yes	Somewhat Yes	Still Unsure	Somewhat No	No
●	●	●	●	●

2. Did you request a Semantic MEDLINE login following the session?

Yes	Plan to but haven't yet	Don't plan to
●	●	●

3. If you don't plan to request a login, could you please say a couple of words on why not?

4. If you did request a login, about how many times have you used Semantic MEDLINE since the Adobe session?

0 times	1-2 times	3-5 times	6+ times
●	●	●	●

5. How likely is it that you will use Semantic MEDLINE for:

	Very Likely	Somewhat Likely	Unsure	Somewhat Unlikely	Very Unlikely
Research	●	●	●	●	●
Facilitating or Finding Collaborations	●	●	●	●	●
Studying/Gaining Knowledge	●	●	●	●	●
Biomedical Literature Exploration	●	●	●	●	●
Hypothesis Generation	●	●	●	●	●

Other (please specify)

6. Do you think Semantic MEDLINE will be a good resource for:

	Yes	No	Maybe	Unsure
Scientists (academic)	●	●	●	●
Scientists (industry)	●	●	●	●
Students (undergraduates)	●	●	●	●
Students (graduate)	●	●	●	●

	Yes	No	Maybe	Unsure
Students (medical)	●	●	●	●
Physicians	●	●	●	●
Other health care professionals	●	●	●	●
Librarians	●	●	●	●

Other (please specify)

7. Will you need additional assistance on the use of Semantic MEDLINE?

Yes	No	Maybe	Unsure
●	●	●	●

If yes, what would you like assistance with?

8. Do you have any additional questions or comments about our instructional session?

9. Do you have any additional questions or comments about Semantic MEDLINE?