PubMed COMMONS: WHAT HAPPENED ON THE WAY TO THE FORUM? RETROSPECTIVE EXPLANATORY CASE STUDY RESEARCH AND LESSONS LEARNED FROM THE U.S. NATIONAL LIBRARY OF MEDICINE'S ONLINE FORUM FOR OPEN SCIENCE

Michelle C. Farabough, B.S., M.S.

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

December 2021

APPROVED:

Barbara Schultz-Jones, Major Professor
Shawne Miksa, Committee Member
Betsy Van der Veer Martens, Committee Member
Jiangping Chen, Chair of the Department of Information Science
Kinshuk, Dean of the College of Information
Victor Prybutok, Dean of the Toulouse Graduate School

Farabough, Michelle C. *PubMed Commons: What Happened on the Way to the Forum? Retrospective Explanatory Case Study Research and Lessons Learned from the U.S. National Library of Medicine's Online Forum for Open Science*. Doctor of Philosophy (Information Science), December 2021, 290 pp., 22 tables, 45 figures, 7 appendices, references, 304 titles.

The U.S. National Library of Medicine brought the intensifying interest in open science to national attention when it joined enthusiastic scientists to introduce and host an Amazon-like rating forum on PubMed—the world's largest database of indexed biomedical and life sciences literature. The result was PubMed Commons. In June 2013, the commenting forum was introduced for open discussion about published scientific literature as part of a three-pronged approach to improve research rigor, reproducibility, and transparency. In Feb. 2018, the forum was unexpectedly discontinued. This retrospective explanatory case study research asked the question, "What happened on the way to the forum?" Answers came from a variety of resources using multiple methodologies for data collection and analysis. Historical data from PubMed Commons' 7,629 comments and 1,551 commenters; key informant interviews with PubMed Commons editors; and a systematized search for published articles, gray literature; and social media content about PubMed Commons were analyzed using computer-mediated discourse analysis and a social network analysis. Results from the quantitative content analysis described a forum with little participation, and the qualitative content analysis demonstrated that active forum members were focused primarily on providing links to other information resources and discussing aspects of post-publication peer review. The social network analysis revealed a disconnected network, which was supported by a sociogram showing a community of independents with only seven small clusters. Findings pointed to 11 factors that affected the forum's adoption and use. Rogers' diffusion of innovation theory scaffolds a forum innovation

agility model developed from this work to offer a better understanding of organizational processes and to aid organizations interested in introducing and managing a similar forum. PubMed Commons was a missed opportunity. No comparable alternative is available to promote open science and serve as a tool for the expected paradigm shift in the way we do scholarly communication in science.

ACKNOWLEDGEMENTS

Writing a dissertation acknowledgement feels much like I imagine preparing an Oscar acceptance speech does. While it’s impossible to mention everyone who has been part of the lengthy, arduous process of completing my Doctor of Philosophy degree, this missive offers an opportunity to publicly declare my heartfelt appreciation to the A-list who has championed and/or tolerated me—whichever the case may be—over the last eight years.

First and foremost, I want to thank my dissertation committee who skillfully guided me with their experience and wisdom. I am forever indebted to my chair, Dr. Barbara Schultz-Jones (aka Dr. SJ), for accepting my request to oversee this final rite of passage in my graduate studies and for granting me unwavering support as I completed the necessary steps to finish writing. She has been a constant sounding board and a believer of my work—always engaged and enthusiastic about my ponderings, explorations, and discoveries. Likewise, I am grateful for Dr. Shawne Miksa, who inspired me with her passion for and knowledge about scholarly communications and the history of information science, and for Dr. Betsy Van der Veer Martens (aka Doc Martens), who has been an unwavering mainstay throughout my graduate studies. I treasure her for consistently meeting me where I am in life and compelling me to be more. Her advice to “keep going” helped immeasurably in the final stretch. All three strong, intelligent women have modeled how to engage students in learning and to stay current in the field.

I am obliged to Shelly Burns, my UNT doctoral program colleague and friend, who introduced me to PubMed Commons and shared my excitement to learn more about the forum’s use and its potential. I’m appreciative of Carolyn Martin for alerting Shelly and I about the NLM’s history of introducing, and then quickly discontinuing programs. This guiding light informed my research and urged me to archive forum data before it disappeared. Similarly, I am thankful for Julia Crawford, my friend and consulting partner, for being co-captain of Team Competent. She has been a lifeline since 2016: actively discussing my work, taking notes during my interview with PubMed Commons editors, entering data into Dedoose, and constantly urging me to finish. Julia's son Xavier and his "data sweatshop" is also owed a debt of gratitude for meticulously entering data into Dedoose.

Other valued colleagues and friends who have traveled this journey with me include: fellow doctoral students Marwah Halwani at UNT and Leslie VanBuskirk at OU; college professors—Drs. Suliman Hawamdeh, Hazem Refai, Cecy Brown, Ryan Bisel, Ana Cleveland, Eric Kramer, Elaine Hsieh, and Young Kim; the Oklahoma Department of Transportation KM Team, especially Michael Molina, April Meadows, Elizabeth Blais, and Trapper Parks; and lifelong soul sisters and cheerleaders—Missy Kruse, Tracy Carroll, Teela Watson, and Michelle Lawyer. I would be remiss not thanking everyone at the Fife Lake, MI library who greeted me with a smile every morning when I showed up to work on the dissertation during the summer of 2021.

More profoundly and personally, I want to thank my sons, Ross and Evan Farabough, and their dad, Don, for their sacrifices so I could read, study, meet, travel, and write. I'm grateful to Ross for inspiring my interest in improving student learning in higher education and to Evan for motivating me by telling me early in this journey that he looked forward to calling me Dr. Mom. Now he can.

I'm extremely proud to be a UNT alumnus.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

It takes little evidence for people to agree that the Internet is at once an amazing innovation and an off-putting disruptor. The shift from traditional to digital media has forced certain changes to our global society, especially as people with connectivity make “going online” a normal part of their day. Advancements in technology have historically come with intended and unintended consequences. While sweeping change benefits some, it troubles others. Online commenting comes to mind—information sharing, social support, and asynchronous connection on one hand; flaming, incivility, anonymity, and misinformation on the other.

Social media sites continue to grow in popularity as people discover their voice is amplified when diffused through the world wide web. Minority groups are uniting online to spread social change, while traditional institutions are becoming democratized. Web-based content is no longer static. Engaging comments and rapid-fire news flashes catch our attention, drowning out messages from familiar media channels and fragmenting our information sources.

Delivering facts and being transparent are more important than ever. Open science is a new approach to a scientific process that’s been the same for centuries. Cooperative work is preferred, and digital technologies diffuse knowledge using collaborative tools. With this movement comes a push for making primary outputs—publications and data—of publicly funded research freely available in digital format (i.e., open access) (Foster Open Science, n.d., para. 1). It has been suggested that such transparency is key for improving research rigor and reproducibility (Collins & Tabak, 2014).

One concern with this reformation is that conventional publishing models are changing. The scientific community is vetting new information communication technologies (ICTs), like blogs, YouTube videos, and online forums, to facilitate scholarly communication. Published letters to the editor, opinions pieces, and expressions of concern have been replaced with widely available and globally instantaneous tweets and online comments.

This dissertation reports retrospective explanatory case study research about PubMed Commons, an innovative alternative to scholarly communication that failed. This online forum was sponsored by the National Institutes of Health (NIH) and open for commenting between its pilot launch on June 13, 2013, and its discontinuation on Feb. 3, 2018. The forum was originally designed as a platform for "open discourse about published articles" (Collins & Tabak, 2014, p. 613). Commenting was publicly visible on the PubMed search engine, and commenters were required to be authors of publications indexed in PubMed.

Case study research is an agile and reflexive means for capturing emergent data and highlighting causal links of a phenomenon too complex to investigate using a single method or data source. This dissertation was designed to shed light on the factors that affected PubMed Commons' organizational and communication processes. An interpretivist approach exposed sensitizing concepts and contextualized social knowledge that can be transferred to other online forums for scientific discussions.

1.1 Statement of Problem

To optimize the adoption and use of new media channels for scientific communication, researchers should investigate how they were developed and how they functioned. Because each type is culturally bound by its purpose and users, investigations focusing on novel digital formats that are intended for foundation-, university-, and government-supported communities and blend professional-reporting with user-generated content and discussion are needed (Bubela, 2009). Despite growing interest in interactive ICTs, there is little understanding of the factors influencing their adoption, how they are being used, and implications for research practices (Proctor et al., 2010). Furthermore, given the abundance of online forums, there has been surprisingly minimal empirical research about their organizational and communication processes (Giles, Stommel, Palus, Lester, & Reed, 2015). This is unfortunate because examining the documentation of naturally occurring discourse and mapping the social connections formed among members of an online forum offers needed insights into these widely used sources of information seeking and sharing. The lack of measurable evidence in this regard is in sharp contrast to the untethered onslaught of new options for online information exchange and the sharing of anecdotal tales. Critical decisions about initiating, maintaining, and managing an online forum are made daily without thoughtful consideration of the lasting, and perhaps costly, consequences of their use. PubMed Commons is a perfect example. The desire to host an online forum can be purely based on assumptions that are fueled by excitement over what can be accomplished using social media technology. While there are obvious benefits to Internet-facilitated scholarly communication, desirable outcomes of establishing and utilizing online forums for scientific discourse should be informed. Organizers could greatly benefit from empirical research that examines existing and failed attempts at online forums intentionally designed to achieve stated objectives.

The NIH brought the intensifying interest in open science to national attention when it joined a group of enthusiastic scientists to host an Amazon-like rating forum for biomedical and life sciences literature on PubMed. The result was PubMed Commons. In June 2013, this online forum was introduced as part of a three-pronged approach to improve research rigor and reproducibility (Collins & Tabak, 2014). In Feb. 2018 the forum was unexpectedly discontinued (National Center for Biotechnology Information [NCBI], 2018). I wondered, "What happened on the way to the forum?"

1.2 Purpose of Study

Scientific practices and scholarly communication benefitting from innovative ICTs are transitioning scientific discovery, collaboration, and information sharing. No longer exclusively paper-based, scientific reporting has evolved to include digital media that lends readers the possibility of accessing information and making timely comments more quickly and easily than previously possible. The purpose of this retrospective explanatory case study research was to examine PubMed Commons and determine how the forum functioned as an online forum and what underlying factors affected its use or nonuse. The goal was providing lessons learned to inform future efforts of organizers who wish to host a similar online forum.

This is important work. Scholars have suggested that communication and information sharing is constitutive of knowledge itself. The fact that PubMed functions as a daily, central access point for a wide spectrum of international biomedical and life sciences researchers and practitioners was a phenomenal opportunity for PubMed Commons stakeholders. "Science benefits greatly from a community that approaches problems in a variety of creative ways" (University of California, 2018, para. 4). Although a linguist, mathematician, communication scientist, and psychologist might each investigate human cognition, each discipline's distinctive approach contributes to a more complete understanding of the phenomenon. The growing emphasis on multi- and inter-disciplinary research underscores the exciting opportunity the PubMed Commons online forum provided PubMed users. Findings from the biomedical and life sciences scientific community is important for every living thing, as this group explores the foundations of life itself, as well as the prevention, control, and treatment of acute conditions and diseases that cause illness and death for humans and animals. Unlike scientists in other areas of science (e.g., astronomy, quantum physics) who seek answers to some of the most fundamental questions about life (e.g., "what's out there?" and "are we alone?"), the biomedical and life sciences community is keen on improving the quality and standard of life here on Earth (Shiode & Parriott, 2016).

The biomedical scientific community is vast, and researchers are dispersed throughout the world. My research aimed to seize the unprecedented opportunity that retrospective explanatory case study research focusing on PubMed Commons affords the next group of pioneering ICT developers and users. By offering a better understanding of information behavior and social communities that formed among researchers who might not otherwise have connected outside of PubMed, I am able to contribute to the expanding field of information sciences. To accomplish this desire, my investigation explored how PubMed Commons functioned as an online forum and what factors influenced the adoption and use of the forum. I was especially interested in the characteristics of commenters, what they discussed, and the types of communities they formed.

1.3 Definition of Terms

A common understanding of terms is necessary to communicate the background, research methods, findings, and conclusions presented in this work.

This work constitutes *retrospective explanatory case study research*: "an in-depth, multifaceted investigation, using qualitative research methods, of a single social phenomenon" (Feagin, Orum, & Sjoberg, 1991, p. 2). The study is retrospective and explanatory in that it looks back on a case that happened in the past and asks "how" and "why" (Yin, 2018, p. 10). To ensure rigorous research, quantitative research methods were also used. Case study research grounds close-at-hand observations and concepts about social action and structures in a natural setting and provides information from a variety of sources over a specified period. This enables a holistic investigation of complex interactions and the complexities of observed social actions and structures (Feagin, Orum, & Sjoberg, 1991, p. 6). Through inductive reasoning and generalization, case studies are particularly important for generating new ideas and theories in the social sciences.

The context for this study has previously been identified as biomedical and life sciences researchers who frequent PubMed (described below), which includes disciplines concerned with understanding, modeling, treating and/or preventing conditions that limit or inhibit life. Scientists in this group are employed in a wide range of professions that include, but certainly are not limited to, geneticists, biologists, pharmacologists, dentists, veterinarians, nurses, physicians and physician assistants, healthcare workers, healthcare administrators, health communication and information specialists, bioinformaticians, and medical librarians/informationists.

This dissertation examined the effect of *Internet-hosted information and communication technologies (ICTs)* among biomedical and life sciences researchers. The Internet is a series of interconnected networks that use standardized communication protocols to facilitate computers worldwide to connect and exchange information. Technological networks are connected in different configurations to form groupings, such as local area networks (LANs) and regional networks. In fact, cell phones are on a network that is considered part of the Internet, as are many other electronic devices. This gives meaning to the term “Internet of Things.” The Internet is distinguished from the World Wide Web, which is the system used to access the Internet. The World Wide Web utilizes HypeText Transfer Protocol (HTTP) to define how messages are formatted and transmitted so that various forms of information available on the world’s different networks can be accessed. One can conceptualize this association as the Internet being composed of the machines, hardware, and data, while the World Wide Web consists of the connections that bring the technology to life.

ICTs refer to the convergence of media technologies and resources used to transmit, store, create, share, or exchange information. These include any device used for communication, including computers; the Internet and its websites, blogs and emails; broadcasting media like radio, television, and webcasting; recorded broadcasting (e.g., podcasting, audio and video players and storage devices); satellite systems; and telephony devices that are fixed or mobile and use visio- or videoconferencing (UNESCO Institute of Statistics, n.d., para. 1).

PubMed serves as the centralized online platform that connects biomedical and life sciences researchers and practitioners by way of the literature it indexes. This free, web-based resource supports the “search and retrieval of biomedical and life sciences literature with the aim of improving health–both globally and personally” (National Center for Biotechnology Information [NCBI], n.d., para. 1). This search engine is maintained by *NCBI* at the *U.S. National*

Library of Medicine (NLM). According to the NLM fact sheet (NLM, 2017),

> PubMed provides free access to MEDLINE—NLM's database of citations and abstracts in the fields of medicine, nursing, dentistry, veterinary medicine, health care systems, and preclinical sciences; citations that provide a record for an article before it is indexed with MeSH and added to MEDLINE or converted to out-of-scope status; citations that precede the date that a journal was selected for MEDLINE indexing; citations to articles that are out-of-scope (e.g., covering plate tectonics or astrophysics) from certain MEDLINE journals primarily general science and general chemistry journals for which the life sciences articles are indexed with MeSH for MEDLINE; citations to some additional life science journals that submit full-text articles to PubMed Central and receive a qualitative review by NLM; and citations for the majority of books and book chapters available on the NCBI Bookshelf. (para. 1)
>
> PubMed also links to full-text articles found in PubMed Central or at publisher websites and other related resources. The database provides advanced search, clinical queries search filters, and special queries pages; links to related articles; and provides discovery tools for other data that may be of interest; includes automatic email for search updates; the ability to save records and enacts filters for search results by way of 'My NCBI'; links to NCBI molecular biology resources; and daily citations.
> (para. 2)

PubMed is a service of the NLM, which is part of the *National Institutes of Health (NIH)*. The NLM is also the world's largest biomedical library.

> NLM carries out its mission of enabling biomedical research, supporting health care and public health, and promoting health behavior by conducting research development on biomedical communications systems, methods, technologies, and networks and information dissemination and utilization among health professionals, patients, and the public. (NIH, 2020, paras. 1, 5, 7)

NCBI is a division of the NLM, established in 1988. Its employees are charged with "creating and maintaining over 40 databases for the medical and scientific communities as well as the general public. NCBI's core literature database is PubMed, which provides abstracts and citations for millions of articles from thousands of biomedical journals" (NIH, 2020, para. 38). "PubMed delivers a publicly available search interface for MEDLINE as well as other NLM resources, making it the premier source for biomedical literature and one of the most widely accessible resources in the world" (Williamson & Minter, 2017, p. 16).

"The NIH provides leadership and direction to programs designed to improve the health of the Nation by conducting and supporting research" (NIH, 2017, para. 3). Its "mission is to seek fundamental knowledge about the nature and behavior of living systems and the application of that knowledge to enhance health, lengthen life, and reduce illness and disability" (NIH, 2017, para. 1). The agency's goals center around innovation, such as fostering "creative discoveries, innovating research strategies and their applications"; developing, maintaining, and renewing "scientific human and physical resources"; expanding the "knowledge base in medical and associated sciences"; and exemplifying and promoting "the highest level of scientific integrity, public accountability, and social responsibility in the conduct of science" (NIH, 2017, para. 2).

The online forum under study was *PubMed Commons*, a virtual space intended for authors to share opinions and information about scientific publications in PubMed (NLM, 2018). The forum was composed of the collection of comments made about publications indexed in PubMed. Comments were publicly visible at the bottom of each abstract on the graphical user interface (GUI) of the PubMed platform. Additionally, the entire collection was originally accessible via a PubMed search for "all[sb]" with the "Reader comments" filter activated. After the forum's discontinuation, data from the forum (i.e., "commented, Pubmedid, Datecreated, [commenter] firstname, [commenter] lastname, [comment] content") was archived on an NCBI file transfer protocol (FTP) site. See full details in Chapter 4.

An *online forum* is an Internet-based webpage that hosts an online exchange of information about a particular topic (i.e., electronic message board). People can interactively and asynchronously post messages, questions, and answers or hold virtual conversations. *GUI* refers to the visual way people interact with an electronic device or computer through its operating system. It enables a person to communicate with the device by way of windows, icons, menus, symbols, and pointing tools. *FTP* websites provide a way for transferring files between computers. This program preceded HTTP for accessing web pages.

A group of individuals interested in a specific topic and who engage in a process of collective learning in a shared domain of human endeavor have come to be known as a *community of practice (CoP)* (Lave & Wenger, 1991). "Online environments are more diverse communities of practice than real-life institutions, being associated with local constructed norms" (Stommel, 2008 as cited in Giles et al., 2015). Many variants and definitions of CoPs have emerged to describe the environment of the community. PubMed Commons' CoP is referred to as an *electronic network of practice (eNoP)*, composed of a typically larger, loose-knit, geographically distributed group of individuals engaged in a human endeavor facilitated through computer mediated communication (Wosak & Faraj, 2005). Members of eNoPs might use email, wikis, or an online forum to exchange information and ideas, as well as to ask and answer questions.

Notably, many members of an online forum are mere observers, yet benefit from the discourse provided by other members who make comments. These participants provide necessary membership in the *latent community* of users (i.e., "a set of interconnected people who share a common interest on a particular subject even when they have not explicitly disclosed their intent of participation or affiliation" (Yoon, Shin, Kim, Par, & Lee, 2009, p. 215). Another popular term for this group is "lurkers."

PubMed Commons leverages Web 2.0 technology—a phenomenon of "the Web's open architecture, its lowering of the barriers to publishing; the ease with which people can connect ideas, the increase in available bandwidth and computing power; …a bottom-up, participatory, rapid innovation, more mixing up and mashing up of information" (Weinberger, 2007, para. 3, 5).

NIH director Francis Collins and principal deputy director Lawrence Tabak (2014) envisioned communication on PubMed Commons *as scientific discourse*. *Scientific* implies the biomedical and life sciences context of PubMed, and *discourse* is distinguished from discussion, in that:

> *Discourse* is a cover term that includes every sort of spoken language, even that used when someone gives an informal talk or presentation to an audience… *Discussion* is a language interaction involving two or more people and carried on in an effort to explore a given topic and perhaps reach some sort of conclusion about it, or at least make progress toward a conclusion. A discussion would therefore also be a type of discourse. (Elgin, 2004, para. 3)

Many comments on PubMed Commons never received a reply, and, even if there was a second comment posted about a particular publication, the latter might not have been related to or in response to the first. For this reason, *threads* are considered a virtual grouping of more than one comment posted on a particular indexed record. In PubMed, these comments appear linear with no indentation. *Nested threads* are those that are hierarchical. In PubMed, initial comments are arranged close to their replies, with a response post generally indented under the original post. Notably, a commenter was not required to select a "reply" radio button to make a comment. Thus, further examination of discussion threads would require *speech act* clues to determine a comment's intention, purpose, or effect.

Because most comments on PubMed Commons were single posts that did not characterize a discussion and because the totality of PubMed users could be considered a CoP, I have adopted the term *Community of Independents,* which was coined by my dissertation committee chair Barbara Schultz-Jones, as a way to describe the community of scientists who were members of PubMed Commons, yet failed to coalesce into a social system of "interrelated units that are engaged in joint problem-solving to accomplish a common goal" (Rogers, 1995, p. 23).

The organizational processes necessary to create a social system to support a sustainable online forum on the PubMed GUI is part of the investigation reported in this dissertation. Many PubMed Commons sponsors, organizers, and users perceived the innovative ICT as a novel medium for *scholarly communication:*

> the system throughout which research and other scholarly writings are created, evaluated for quality, disseminated to the scholarly community, and preserved for future use. The system includes both formal means of communication, such as publications in peer-reviewed journals, and informal channels, such as electronic listservs. (Association of College & Research Libraries [ACRL], 2003, para. 1)

Accordingly, scientific discussions on PubMed Commons could be considered *informal* scholarly communication (i.e., discourse to communicate aspects of scientific work that occurs in the absence of formal social structures). Comments were not formally requested, nor were they governed by traditional rules of publishing conduct, like *peer review.* Although PubMed Commons comments were prompted by *formal* scholarly communication, which includes physical and digital publications, commentaries, opinion pieces, letters to the editor, conference proceedings, and formal presentations, they were not preserved as part of the scholarly record. This presented a problem, as commenters intended their posts as an alternative to or replacement for formal scholarly communication. In fact, an evolving publishing model that favors open science is drawing attention to this concept of post-publication peer review (PPPR).

Traditional research output is typically approved through *peer review*—the process by which formal scholarly communication is evaluated by others who are practicing members of a particular field and mediated by an editor or editorial board. *PPPR* is peer review that occurs after an article is published. Conceived as part of open science, PPPR advocates argue that open evaluation with transparent peer review and paper ratings will increase the overall quality of the peer review process and promote self-corrective science. Wider availability and growing popularity of Web 2.0 platforms enable rapid assessment in real time and typically via social media. Insights into the impact and influence of research have until recently been measured by way of traditional bibliometrics (e.g., citation count and journal impact factor), which take time to accrue. Technology-base, alternative-level metrics (i.e., altmetrics) offer new ways to measure effectiveness by considering the attention and impact of scientific publications, even if the influence is more short-lived (Rong, Lopes, Hameed, Gaudino, & Charlso, 2020). The effect is nonetheless valuable for advancing science. With these changes in mind, talk of a publishing paradigm shift in scientific practices is rapidly increasing, fueling emotions and grabbing scientists' attention (Knoepfler, 2015).

One condition that regularly sparks debate related to online posting is *anonymity* (i.e., the condition of being anonymous). Commenters on PubMed Commons were required to use their real name when posting. Possible effects on forum adoption are discussed in Chapters 4 and 5.

One motivation for this dissertation was the way in which the Internet acts as a major disruptor to traditional scientific communication. Lines between informal and formal scholarly communication are now blurred. In 2013, PubMed Commons joined PubPeer, Retraction Watch, F1000, and arXiv as a platform where PPPR can readily take place (Knoepfler, 2015). One point of contention is what constitutes a peer. Other prickly issues encompass *publishing models* (i.e., calculated business approaches to financing and delivering content to users) and media fragmentation. These are discussed in Chapter 5.

Information behavior has been described as "the totality of human behavior in relation to sources and channels of information, including both active and passive information seeking, and information use" (Wilson, 2000, p. 49). Patterns of commenter information behavior (e.g., how often a comment was posted; whether the comment was made in isolation or as part of a thread), are detailed in Chapter 4.

The research methods of my study included computer-mediated discourse analysis (CMDA) as a scientific approach to studying computer-mediated communication (CMC), in other words, communication that occurred through networked computers. *Message thoughts* were used as the unit of analysis for examining PubMed Commons comments. In this way, a single post could be coded as multiple message thoughts (i.e., categories). *Archival data* (sometimes referred to as historical data) is information that already exists in files or documents. In this dissertation, these include demographic data to characterize commenters, as well as a downloadable spreedsheet containing all comments posted on PubMed Commons.

Key informants are individuals who are interviewed about a particular organization, problem, or research interest. PubMed Commons editors served as my key informants.

Recently, there has been a growing interest in *social network analysis (SNA)*. SNA is an approach and a set of techniques applied to the study of the invisible relational structure of groups (Schultz-Jones, 2009, p. 595). SNA was used to examine the extent to which eCoPs formed on PubMed Commons. SNA has previously been used in the biomedical and life sciences to investigate phenomenon like epidemics and communicable diseases. Citation analysis can be considered SNA; information scientists have used it in this way to map the diffusion of scholarly communication.

1.4 Research Questions

The overarching research question for this dissertation (i.e., RQ 1) asked, “how did PubMed Commons function as an online forum for scientific discourse?” This question was further refined into :

- RQ 1a—how often and in what ways did stakeholders participate in PubMed Commons?
- RQ 1b—what were characteristics of forum commenters on PubMed Commons?
- RQ 1c—what was the subject matter of comments posted on PubMed Commons?
- RQ 1d—what types of communities formed on PubMed Commons?

A follow up question (i.e., RQ 2) asked what factors contributed to or limited the adoption and use of PubMed Commons? The holistic investigation of the forum and the complexities of observed social actions and structures resulted in the development of the forum innovation agility model, which is highlighted in Chapter 5.

1.5 Study Significance

Participating in an online forum is a daily event for many people who are privileged with access to the Internet. Those who aren't participating in a forum are likely latent users who are benefiting from the information exchange. I discovered there is limited research about organizational and communication processes of online forums, as well as the information behavior associated with a large-scale, open access online forum targeted at scientists, especially one that has been integrated into a government-sponsored ICT platform, like PubMed. Several government agencies have shared their interest in this type of technology, and my work could inform them as they contemplate introducing and managing a similar forum. The forum innovation agility model I developed highlights factors that should be considered. It also suggests measurements that might be employed and features that could quickly be adjusted to mitigate forum adoption and ensure sustainable use.

Methods for capturing data, quantifying influencing factors, and reporting rich, thick descriptions about online forums is complex and underreported. The mixed method approach modeled in this retrospective explanatory case study research and its application for studying PubMed Commons could help future researchers design rigorous research of other online forums.

1.6 Assumptions

When I first began studying PubMed Commons in 2014, I naively made an underlying assumption that PubMed Commons would be successful in hosting scientific discourse and that there would be sufficient participation to justify the forum's sustainability. After visiting with PubMed Commons editors, this assumption was perpetuated. I continued to assume that the forum would remain open, regardless of its adoption or use. At the same time, I also assumed that the targeted community of biomedical and life sciences researchers was too large and diverse to coalesce into single eCoP. I anticipated that the NLM would articulate the definition of “peer,” clarify the purpose of the forum, and actively promote ot use. Realistically, I presumed that the strong opinions and emotions about changes in scholarly communications (i.e., PPPR) would loom over the forum and affect comment content and forum adoption. Following the discontinuation of PubMed Commons, I assumed the decision to cease PubMed Commons was premature and that an ample amount of time was not given for strategic promotions to build the eCOP, to achieve a realistic rate of adoption, and to increase the number of comments. I also assumed that there were mitigating factors affecting the forum’s organizational processes that my research would uncover. Based on my experience as a knowledge and project management consultant, I presumed that a) the forum was not effectively promoted; b) required registration and full disclosure of commenter identity affected adoption; and c) rapid changes in publishing models and scholarly communication were causing a paradigm shift in scientific research practices.

More than anything else, I assumed that like me, targeted users were confused about the purpose of PubMed Commons and questioned what type of comments they should share.

1.7 Chapter 1 Summary

Retrospective explanatory case study research examining the PubMed Commons online forum is urgent, as an increasing number of federal agencies are mandating public access to government-funded research. Several agencies are considering the use of ICTs to connect scientists and citizens. In spite of this, forums hosting after publication commenting continue to close (McKenzie, 2017; Spector, 2013). This phenomenon is in sharp contrast to the fact that an increasing number of people rely on ICTs for connecting them with information resources and with each other. I designed "how" and "why" research questions consistent with retrospective explanatory case study research to investigate how PubMed Commons functioned as an online forum for open discussion about published articles. I also investigated what factors might have affected a low adoption rate, which ultimately led to the forum's discontinuation. Learning about the characteristics and information behavior of people who made comments, the content of their comments, and whether they formed eCoPs was important for developing a theory about what might have happened on the way to the forum.

This chapter introduced readers to the background of PubMed Commons and why studying the online forum was important. I explained that scholars believe communication and information sharing is constitutive of knowledge and that given widespread use of the internet, there should be additional research about the way scientists use ICTs to communicate about their work. Terms used throughout this dissertation were defined to ensure mutual understanding as a reader progresses from a literature review and explanation of methodologies to the reporting of results and a discussion about my findings. This section also provided additional details about the context of this study.

Research questions were articulated, and I shared two significant outcomes of my work: 1) extending knowledge about the organizational and communication processes for introducing and managing an online forum and 2) understanding how the methodology used in this case study could advance the way information scientists investigate online forums.

The following chapter reports a literature review that gives more details about the history of online forums and the NLM's role in innovating technologies for enhancing access to health services research information, toxicologic and environmental health data, and clinical trial information. Chapter 3 explains my methods for data collection and analysis of PubMed Commons. Answers to my research questions are thoroughly explained in Chapter 4 with results presented in figures and tables. A final chapter uses Everett Rogers' Diffusion of Innovation Theory to explain 11 factors I believe affected the adoption and use of PubMed Commons. Based on these, I developed an ecological model, namely the forum innovation agility model, that organizes these factors into layers on a spectrum ranging from elements inside the control of forum innovators and organizers to those outside their control.

CHAPTER 2

LITERATURE REVIEW

Chapter 2 of this dissertation highlights literature necessary to better understand the nature and purpose of this study. A methodological review of past literature is an essential enterprise for any research. Researchers are charged with uncovering what has previously been discovered to build on a corpus of knowledge and strengthen a field of study. A meaningful literature review aids in achieving this mission by using "ideas in the literature to justify the particular approach to the topic, the selection of methods, and [to demonstrate that the] research contributes something new" (Hart, 1998, p. 1). To be effective, the literature review in a dissertation should be appropriate in depth and scope; demonstrate rigor; efficiently analyze and synthesize published findings; and clearly narrate the current body of knowledge known about the topic under investigation. To accomplish these goals, this chapter is divided into three main sections, which are briefly summarized in the paragraphs below, and then fully detailed in the pages that follow.

The first section of this chapter provides pertinent background information about the U.S. National Library of Medicine (NLM), its National Center for Biotechnology Information (NCBI), and the PubMed/MEDLINE database search engine which hosted the PubMed Commons online forum. In the second section, the phenomenon of online forums and their communities is explored, and with this, a review of past and current research to provide a better understanding about forums as a type of information communication technology (ICT). The third section informs about scholarly communication as it transitions from a centuries old process to one that includes ICTs and a push for open science. Knowledge and information sharing among researchers, such as the comments that constituted PubMed Commons, is widely recognized as form of scholarly communication. A distinction between formal and informal scholarly communication is articulated, along with Insights about the ways in which the Internet is blurring the line between these two types. Special emphasis is given to the peer-review process, especially post-publication peer review (PPPR) and the debatable role it is expected to play in ensuring research rigor, reproducibility, and transparency.

2.1 U.S. National Library of Medicine

In 2016, the NLM celebrated 180 years of advancing biomedical and life sciences knowledge. “From its modest beginning in 1836 as a shelf with a handful of medical books that constituted the Library of the Office of the Surgeon General of the Army, the NLM has grown into the world’s largest biomedical library” (Slomski, 2011, p. 2158). The agency was reassigned from the Armed Forces to the U.S. Public Health Service and officially rebranded the NLM in 1956. Former NLM director Donald Lindberg (2011) reported that the library’s purpose has always been laser clear: “to acquire, organize, disseminate, and preserve the biomedical knowledge of the world to promote scientific advance and public health” (p. 46). Today, in addition to indexing a collection of more than 32 million journals, manuscripts, audiovisuals, newsletters, online books, and other materials in more than 150 languages, NLM-maintained databases and electronic tools deliver trillions of bytes of scientifically based digital information and data to worldwide users.

Recognized as part of the U.S. National Institutes of Health (NIH) since 1968, the NLM is funded through the U.S. Congress and is tasked with making biomedical and life sciences resources freely available. Software technologies have been designed specifically to facilitate the organization, searching, and retrieval of scientific research data/reporting, including molecular biology and genomic information (National Institutes of Health [NIH], 2017). The NLM's web sites and the over 40 databases it manages are freely accessed several million times each day by individuals with an Internet connection (Lindberg, 2011).

Congressional actions have charged the NLM with special responsibilities for enhancing access to health services research information, toxicologic and environmental health data, and clinical trial information (Lindberg, 2011, p. 47). Since 1982, the GenBank database of nucleic acid sequences has supported biomedical and biological research around the globe. The Visible Human Project was unveiled in the 1990s and features a digital library of images representing the complete anatomy of both a male and female. In addition to its value for studying anatomy, the data sets are used "for a variety of medical, scientific, and nonmedical purposes—from practicing surgeries and medical procedures to designing furniture and machinery that is comfortable to use" (Lindberg, 2011, p. 47).

The NLM introduced the MedlinePlus.gov consumer health website and database in 1998, providing patients and families across the globe with health information about hundreds of health-related topics. The ClinicalTrials.gov database was introduced in 2000 and serves as an online registry of new and ongoing clinical trials. That same year, the PubMed Central database was launched to provide free full-text digital access to nearly 5 million biomedical and life sciences journal articles.

The NLM has a well-established, proven record of accomplishment for providing access to a variety of innovative and useful information products and resources. This demonstrates that the NIH values information sharing and seeks ways to leverage new technologies for engaging researchers and the public with biomedical data and health information. In addition to the aforementioned successes, the NLM has also introduced experimental services, like the now discontinued PubMed Commons online forum. An early example of the NLM's foray into scholarly communication came in the early 1960s when biological preprints were circulated via Information Exchange Groups (IEG). "Although the system attracted over 3,600 participants and saw the production of over 2,500 different documents, by 1967 it was effectively shut down following the refusal of journals to accept articles that had been circulated as preprints" (Cobb, 2017, p. 1).

The NCBI is a division of the NLM that was established in 1988. Its employees are charged with "creating and maintaining over 40 databases for the medical and scientific communities, as well as the public. NCBI's core literature database is PubMed, which provides abstracts and citations for millions of articles from thousands of biomedical journals" (NIH, 2020, para. 38). Comments on PubMed Commons appeared below abstracts on the PubMed graphical user interface (GUI). "PubMed delivers a publicly available search interface for MEDLINE as well as other NLM resources, making it the premier source for biomedical literature and one of the most widely accessible resources in the world" (Williamson & Minter, 2017, p. 16).

2.1.1 PubMed/MEDLINE Database

The PubMed/MEDLINE database is likely the best known and most heavily used of the NLM's myriad of electronic resources. The database's underpinnings can be traced to 1879 when John Shaw Billings, M.D.—appointed to supervise the Surgeon General's Library—indexed, catalogued, and then published his bibliographic records in the first volume of the

Index Medicus. The publication grew exponentially and with the help of new technologies went out of circulation in 2004 “due to the impossibility of managing voluminous information generated in the form of a book” (Bravo, 2016, p. 5). Migration to online content began in the mid 1960s and continued into the 1970s with earlier digital versions of MEDLINE. The transition began with the Medical Literature Analysis and Retrieval System (MEDLARS). When first implemented, MEDLARS “was heralded as the first library application of a computer to handle scientific literature with a digital computer and high-quality composing equipment” (Dee, 2007, p. 419). Based on recommendations from the NLM’s Index Mechanization Project (July 1958 to June 1960), the system aimed to improve publication indexing and create a by-product bibliographic retrieval system. The first phase of MEDLARS became functional in 1964 with outputs of printing products, recurring bibliographies, and demand (i.e., nonrecurring) bibliographic searches. To keep pace with new technological developments, MEDLARS II leveraged the IBM 360 series computer system, an automated library system with an automated acquisition and cataloging system, an improved indexing and search aid (i.e., online Medical Subject Headings [MeSH]), an ability to provide chemical compound and toxicological searches, and a high-performance graphic image storage and retrieval system. The upgraded system was complete on January 3, 1975, and a new NLM five-year strategic plan initiated the development of an online version, namely “MEDLINE,” referring to MEDLARS online.

An electronic version of MEDLINE was first introduced in January 1996, as an “experimental” database under the Enterez retrieval system. The public interface for the new PubMed search engine website was officially introduced the following year at a Capitol Hill ceremony led by then Vice President Al Gore and ranking Labor/Health and Human Services

Appropriation Subcommittee Members—Senators Tom Harkin (D-IA) and Arlen Specter (R-PA) (Smith, 2013). About PubMed, Gore announced the information resource "may do more to reform and improve the quality of health care in the United States than anything we have done in a long time" (Slomski, 2011, p. 2158). With this declaration, the world suddenly had access to the entire MEDLINE bibliographic database.

MEDLINE remains the primary component of PubMed, containing more than 28 million references (NLM, 2021) to biomedical and life sciences journal articles that are indexed with NLM MeSH to assist users in searching and retrieving its scientific information. This feature sets MEDLINE metadata apart from the rest of indexed citations in PubMed. Journals for inclusion as part of MEDLINE are recommended by the Literature Selection Technical Review Committee, which is an NIH-commissioned advisory committee of external subject experts, much like committees that are appointed to review the institute's grant applications. Additional journal titles and newsletters focusing on special priorities for the NIH/NLM (e.g., health services research, toxicology and environmental health, molecular biology, complementary and alternative medicine) are selected following internally initiated reviews that are conducted by external reviewers and organization, as well as NIH experts.

Currently, citations from more than 5,200 scholarly journals (NLM, 2021) published around the world in at least 40 languages are generated by the NLM as part of the MEDLINE database records. International partners (e.g., African Journal Partnership Project Program, Karolinska Institute, PubMed Central International) and collaborating organizations (e.g., International Health Terminology Standards Development Organisation) are involved in creating and adding citations to MEDLINE each day. Nearly 952,919 citations were added in 2020 (NLM, 2020).

2.1.2 Digital Formats for Information Dissemination

The transformation of scholarly communication to digital and electronic formats of articles and the increased public use of the Internet in the early 1990s changed the way scientific research is disseminated. This new concept for converting biomedical libraries from centers of collection and organization for printed medical literature and other physical information resources to hubs for the management and administration of information and medical knowledge was also revolutionary. Subsequent changes at the NLM have been significant.

2.1.2.1 PubMed Search Engine

The freely accessible PubMed search engine currently provides access to nearly 32 million indexed records for biomedical and life sciences abstracts and citations of journal articles, proceedings, and online books (NLM, n.d.). Anyone with an Internet connection can search the database for hyperlinks to full texts, many of which are open access. Some require authentication through publisher subscriptions paid by a researcher's home institution.

In addition to the MEDLINE-indexed subset of journal records, citations retrievable in the PubMed search engine also include journals and manuscripts deposited in PubMed Central (PMC)—the NLM's digital collection of open access (i.e., free) journal articles—and literature indexed in the NCBI Bookshelf. Other records include citations and hyperlinks for: a) in-press articles before they are indexed with MeSH; b) out-of-scope articles (e.g., plate tectonics or astrophysics) from certain MEDLINE journals; c) “ahead of print” records that precede an article’s final publication in a MEDLINE-indexed journal; d) articles that precede the date a journal was selected for MEDLINE-indexing; e) pre-1966 articles that have not been updated with current MeSH status; f) articles from additional life sciences journals that submit full text to PMC and receive qualitative review by the NLM; and g) manuscripts of articles published by NIH-funded researchers (NLM, 2017c).

On February 15, 2018—the day PubMed Commons was discontinued—over 28 million PubMed records were listed with their abstract, and 18.3 million had links to full-text, of which over 6 million were available for free (i.e., open access). Interested parties can obtain information about the current size of the database by typing “all[sb]” into the PubMed search bar at the following url (https://www.ncbi.nlm.nih.gov/pubmed/), and then clicking “search.” A similar search on Sept. 11, 2021, retrieved 33,052,810 records. Once results are displayed, a user can then select “abstract,” “free full text,” or “full text” format from the left column of available filters. What makes PubMed and other NLM information resources unique is that scientists, health professionals, and the global public alike have access to the same database records, making the information shown on the GUI fully transparent to everyone.

2.1.2.2 PubMed Commons

In 2013, the NIH launched PubMed Commons—an online forum with membership aimed at the authors of literature indexed in the PubMed search engine database. The purpose was “open discourse about published articles” (Collins & Tabak, 2014, p. 613). The targeted audience for the forum is highly regarded in that “the nation’s academic biomedical research community provides essential services that underpin American society” (National Academies of

Sciences, Engineering, and Medicine, 2017, para 5).

The first PubMed Commons comment was posted on June 12, 2013, at 9:51 p.m. by Robert Tibshirani, a Stanford University professor in the Departments of Statistics and Health Research and Policy and member of the Meta-Research Innovation Center at Stanford (METRICS, n.d.) team. Tibshirani has been credited as an initiator of PubMed Commons and the leader of a group of scientists who worked with then NCBI director David Lipman to make the envisioned commenting system on PubMed a reality. The new online forum was intended to be a place where scientists could exchange ideas, ask questions about methods/techniques, offer suggestions, make comments on each other's work, and even offer praise (Spector, 2013, para. 7). See Chapter 5 for a more in-depth history of PubMed Commons.

Tibshirani's first PubMed Commons post set the tone for the forum. In his comment, which appeared under the abstract to an article entitled "Detecting Novel Associations in Large Data Sets," he referenced comments he published in *Science*, which questioned the authors'/researchers' methods and suggested the use of distance correlation as a preferred method of measure. The nature of this post was in direct accordance with the NIH's aim to utilize PubMed Commons for open discourse about published scientific articles.

PubMed Commons was distinct from similar online forums in that it provided both a commenting platform and an enormous global audience that regularly visited the PubMed website. "On the average working day approximately 2.5 million users from around the world access PubMed to perform about 3 million searches and 9 million page views" (Fiorini, Lipman, & Lu, 2017, para 2).

PubMed Commons was just one part of a three-pronged approach introduced in 2014 to improve research reproducibility. Additional initiatives included a) developing and incorporating a training module on enhancing research reproducibility and transparency—with emphasis on sound experimental design—into the mandatory training on responsible conduct of research for NIH intramural postdoctoral fellows (see https://www.nih.gov/research-training/rigor-reproducibility); and b) enhancing the NIH Big Data initiative by developing a data discovery index (DDI) that allowed investigators to search and access unpublished, primary data (see https://datascience.nih.gov/bd2k/funded-programs/resource-indexing). The NIH also introduced a checklist to ensure a more systematic evaluation of grant applications by reviewing experimental design features (see: https://grants.nih.gov/reproducibility/index.htm).

The newly established PubMed Commons Blog heralded that its purpose was to "enable authors to share opinions and information about scientific publications in PubMed" (see Appendix A). The blog's first post welcomed users on Nov. 14, 2013, and announced that on Nov. 26, 2013, the first version of the experimental pilot project would be publicly available. A hyperlink connected potential users to instructions about joining PubMed Commons and to a report of early developments in the PubMed Commons pilot. In addition to the blog (https://pubmedcommonsblog.ncbi.nlm.nih.gov/—no longer accessible, but archived in Appendix A), the "PubMed Commons team" introduced a PubMed Commons Twitter account (https://twitter.com/pubmedcommons/—no longer available). Besides occasional PubMed Commons Blog posts and regular tweets, major developments or updates to PubMed Commons were announced on the NCBI Insights blog, Facebook page, Twitter account, and YouTube channel.

Guidelines for PubMed Commons were accessible on an NCBI-maintained webpage until the forum's discontinuation (see Appendix B). The rules of engagement were clear about what constituted unacceptable activity that would be removed (Couchman, 2014, p. 9). Members of the forum (i.e., a term used by the PubMed Commons team to describe its commenters) were required to have an NCBI account, as well as an invitation to join the forum. The latter was accomplished through self-selection from a pre-approved list of email addresses of individuals identified as eRA Commons members, NIH intramural researchers, or persons funded by the Wellcome Trust. Alternatively, any author with an article indexed in PubMed could receive an invitation from someone who was already a member of PubMed Commons, after supplying them with the PubMed Identification (PMID) of their publication record. Groups of 50 or more PubMed authors could also send their names and e-mail addresses to NCBI organizers. Membership in the forum was extended to Journal Clubs on Dec. 17, 2014. This group was asked to "share key points, questions, and summaries from their discussions'' (see Appendix A).

Although authenticated commenter names were publicly available on the PubMed GUI, a system for commenter "profiles" was not established. Ratings of comments by members formed the basis of scores for commenters, as well as individual comments. Scoring affected the display of all comments made by an individual commenter (NLM, 2017b). The more an individual participated in PubMed Commons, the higher their score and potentially more recognition among forum users. Deleted comments also affected individual commenter scores.

There is much speculation about the decision to discontinue the PubMed Commons online forum. Comments posted in reaction to the announcement (e.g., NCBI Insights announcement, tweets, and blog postings around the globe) were part of the systematized search for publications, gray literature, and social media content used as evidence to support claims reported in Chapter 5 of this dissertation. Some people suggested guidelines deterred commenters. As mentioned above, forum members had to establish an NCBI account and use their real identity when commenting (i.e., no pseudonyms or anonymous accounts were allowed). Members were also required to disclose potential conflicts of interest and not use PubMed Commons to systematically promote a product, position, or their own publications. Members were not allowed to target other publications; share partisan political views; plagiarize content; allege misconduct of authors, reviewers, editors, and publishers; speculate about the motivations of authors, reviewers, editors, and publishers; or use discriminatory, racist, offensive, unlawful, or derogatory language (see Appendix B). An overseeing committee found it necessary to add the term "inflammatory" to the guidelines in 2015 (H. Bastian, personal communication, October 28, 2016).

Submitted comments were initially filtered through an automated check for inflammatory language and then posted immediately. Posts were reviewed online and those that violated guidelines were temporarily held offline for review by moderators. Given that a comment was posted after the initial automated screening and that a moderator later found it necessary to remove the comment, the post was replaced on the PubMed GUI with the date of comment, name of commenter, and a notation that the comment was removed by moderators. A member who post violated a guideline was contacted and invited to revise/resubmit his/her comment, which many did, as evidenced by a comment marked as "removed by moderators" on the PubMed BUI followed by another full comment by the same commenter (H. Bastian, personal communication, October 28, 2016). In some cases, the follow-up post showed that the comment had been "edited" or "deleted" by the commenter. The former would indicate that the commenter and comment were working toward compliance with guidelines; the latter would indicate refusing to comply with guidelines.

While the forum was operating, comments appeared beneath abstracts in reverse chronological order on the PubMed GUI. Replies to comments were nested under the original comment in the order in which they were posted. Ratings for individual posts automatically determined which comments appeared as "top comments now" on the PubMed Commons Blog homepage. This information also appeared on the PubMed search engine GUI.

Members were invited to report a concern or lodge a complaint about a comment. Guidelines stated that comments were intended as a response to the content of a publication, not as a method for reporting typographical errors, duplicate records, or broken links. Comments were initially limited to 8,000 characters; however, PubMed Commons team members quickly noted that commenters got around this rule by posting a succession of comments. In response, the limit was removed (H. Bastian, personal communication, October 28, 2016). Many comments served as alternatives to formal scholarly communication (e.g., letter to the editor, brief communications, editorial, correspondence). Providing hyperlinks to a referenced article or additional information on external sites (e.g., Twitter, personal blogs, data sets, other publications) was encouraged. This practice of redirecting discussions to alternative media is addressed in Chapter 5.

Although comments could be edited or deleted, commenters were told the permalink to their comment would remain valid as a form of documented scholarly communication. Unfortunately, these are no longer available. To date, the official record of comments is available from an NCBI FTP website in a "commons_archive.csv" file (

https://www.ncbi.nlm.nih.gov/pubmedcommons/). By default, commenters agreed to grant PubMed Commons a non-exclusive, irrevocable, royalty-free license—under the Creative Commons Attribution 3.0 License—to distribute his or her comment to the rest of the world.

PubMed Commons received limited mass media attention or recognition in the published literature. Blog posts and journal editorials focused their attention on PPPR (see examples in Chapter 5). Designation of "peer" was questioned.

> This [PubMed Commons] enables greater interaction and communication within the scientific world and allows for rapid proliferation of knowledge from one end to the other… The greatest achievement of PubMed Commons is to expedite post-publication review and transfer of knowledge. This is a faster way to share as opposed to the current status in which these communications might take 3 to 6 months to get available for the readers… A negative comment by an inexperienced reader might deter others from using potential valuable material. (Hasan, Masood, & Memtaz, 2016, p. 913)

> The question is whether added comments will be useful or subject to misuse, and indeed, if this commentary is "peer" reviewed. Will comments be applied to papers by contributors with real expertise in the area? In addition, will authors of papers that are listed in PubMed check regularly to see who is attaching comments and respond? (Couchman, 2014, p. 9)

A *Neuron* editorial written by authors who had received criticisms in several venues (e.g., social media, on *Neuron's* own commenting forum, and in PubMed Commons) encouraged a constructive debate of their published papers.

> …the exchange of ideas is the fuel of science. Providing a robust, peer-reviewed evaluation of the arguments raised stands as an important complement to other forms of commentary. We hope that by providing a curated platform for communicating these peer reviewed critiques and responses in the journal we are contributing to an open and constructive scientific discourse. It is our belief that the community will benefit from witnessing this scientific exchange and that the process of science examining itself will naturally lead to progress. (Wang et al., 2016, p. 331-332)

As PubMed Commons matured, people who were watching, researching, or using the forum voiced their concerns about low usage and disappointing participation.

> If PubMed Commons is to fulfill its proposed role in the post-publication evaluation of scientific research, levels of adoption must improve, and commenters must disclose any pertinent conflicts of interest. (Lane, 2016a)

A pilot study I completed with my colleague Shelly Burns that investigated forum activity between October 2013 and July 2015 found that less than 0.5 percent of eligible records had received a comment (Farabough & Burns, 2015). This figure was validated by two similar studies (Lane, 2016; Ramos et al. 2015). The qualitative content analysis of our study (n=232/2,500 or 9%) identified 11 major themes repeated in single and threaded comments. These included "watchdog, disputing, redirecting, discussing, public forum, flaming, validating, humor, inquiring, promoting, and author response." Lane (2016) found similar content, and specifically mentioned that comments raising concerns about an article focused on "omission of important citations, CONSORT issues, ClinicalTrials.gov issues (lack of NCT registry or posted results), plagiarism, and retractions" (para. 25).

Lane's poster presentation at the International Society for Medical Publication Professionals also reported that the highest number of comments on any given article was 17 and that most articles received less than two comments. Many articles with comments were openly accessible and/or published in top-tier journals; they displayed scientific rigor or fulfilled a scientifically useful purpose. The typical comment (22.4%) was a response with citations and resembled a letter to the editor. Second most common was a more general comment (16.7%), followed by a comment that linked to supporting research (8.2%), conflicting research (6.0%), or a blog post (7.1%). Links to free full text, supplementary data sets, or data analysis software were also found.

At the same conference, Ramos et al. (2015) reported that despite fluctuations in the rate of commenting during the early pilot phase of PubMed Commons, there was no evidence of an increasing trend. The group also found that 30% of comments about clinical and pharmacological articles were negative and 17% were positive. The greatest number of overall comments (52%) were neutral. Only 1% of the comments were removed by moderators. Number of comments and journal impact factor did not correlate, although there was significant positive association between articles in high-impact general medical journals and number of comments. The researchers suggested that allowing a wider audience to comment might increase its utility.

The PubMed Commons team enacted mitigating measures to increase forum adoption. Noting that many comments posted on PubMed Commons were not receiving responses from authors, a system to contact authors when a comment was added to their PubMed record was added in 2015. Although PubMed Commons lead editor Hilda Bastian assured me that activity on the forum was not a concern (H. Bastian, personal communication, October 28, 2016), a February 2018 blog post on NCBI Insights indicated that low activity led to the decision to discontinue the forum: “While many worthwhile comments were made through the service during its 4 years of operation, NIH has decided that the low level of participation does not warrant continued investment in the project, particularly given the availability of other commenting venues” (NCBI, 2018, para. 4).

2.1.2.3 PubMed Commons Potential

The PubMed search engine has been identified as a critical asset to humankind particularly because it facilitates information retrieval of health research about risk factors and trends in diseases, protocols for care, treatment outcomes, public health interventions, and health care costs and use (Institute of Medicine, 2009). Its audience is unique in that they are focused on the foundations of life itself and the prevention, control, and treatment of acute conditions and diseases that cause illness and death for humans and animals. “The nation’s academic biomedical research community provides essential services that underpin American society” (National Academies of Sciences, Engineering, and Medicine, 2017, para. 5).

The larger PubMed community is employed in a wide range of professions that include, but certainly are not limited to, geneticists, biologists, pharmacologists, dentists, veterinarians, nurses, physicians and physician assistants, healthcare workers, healthcare administrators, health communication and information specialists, bioinformaticians, and medical librarians/informationists. These professionals are dispersed throughout the globe, making the potential size of the PubMed Commons community immense; in fact, its potential is unmatched. Furthermore, many of the individuals who use PubMed might not otherwise come into contact outside their searching behavior on this information resource. PubMed Commons seemed well positioned for success, as the PubMed GUI provided both an audience of potential commenters and a centralized nucleus for scientific literature and its authors. Unfortunately, forum adoption increased at a slow rate, and commenting started to decline in 2015.

2.2 Online Forums

Since its beginnings in the late 1960s, the Internet has supported rapid developments in ICT and an accompanying increase in computer-mediated communication (CMC). People and information have been connected in new ways. Email, social media, interactive websites, wikis, and real-time news empower users to create content and exchange information at will. Communication can be either instantaneous or asynchronous.

Table 2.1

Advantages and Disadvantages of Online Forums

Advantages	Disadvantages
Flexibility—Given connectivity, platform is accessible anytime, anywhere	Text-based—Currently technology generally relies on inputting text, which can be challenging for those who don't like to write or have poor keyboard skills. With the advance of broadband connectivity and voice and video conference technology, this will be less of an issue.
Leveling—Reserved people who infrequently "speak up" are not drowned out or interrupted by "louder" others can't interrupt	No physical cues—Without facial expressions and gestures or the ability to retract immediately, there's a big risk of misunderstanding. Again, advances in technology could change this.
Documented—Unlike verbal conversation, online discussion is lasting and can be revisited	Information overload—A large volume of messages can be overwhelming and hard to follow, even stress-inducing.
Encourages reflection—Participants are not required to contribute until they've thought about the issue and feel ready.	Threads—The logical sequence of discussion is often broken by users not sticking to the topic (thread).
Relevance—Provides a place for real life examples and experience to be exchanged.	Time lag—Even if a person logs on daily, 24 hours can seem like a long time when waiting for a reply; by then the discussion could have changed course or moved on.
Community—Over time connections can develop into a supportive, stimulating community that participants value.	Inefficient—The interaction takes longer than verbal conversation, thus it's hard to reply to all the points in a message. Questions are often unanswered.
Limitless—Discussions are free to flow; the unexpected often results in increased incidental learning.	Isolation—Some individuals prefer not to participate in online discussions, thus are left out of the conversation.
Choice—Commenters have a choice of contributing either a quick question/observation or a long reflective contribution.	Directionless—Commenters used to having direction can find a leader-less environment overwhelming.

Technologies like online forums have become a common hub where people with common interests gather, free from limitations imposed by geography, time zone, and cultural background (Zhang, 2007, p. 351). Virtual communities enable members to contribute content and leave with new knowledge (Constant, Sproull, & Kiesler, 1996; Wasko & Faraj, 2005; Zhang

& Watts, 2004). Community members can instantly contest or correct information. Online interactions sometimes develop into real-world socializing or, on the other hand, connect groups who have previously gathered at face-to-face conferences. In the wake of COVID-19, these types of gatherings are more frequent and increasingly important. Table 2.1 summarizes advantages and disadvantages of online communication.

2.2.1 Historical Foundations of Online Forums

Online forums were born in the spirit of community. Before the Internet developed into the current mass communication medium we enjoy today, pioneering computer scientists conversed and exchanged files using “bulletin-board systems” (BBS). The first recorded use of a home computer for hosting messages was the Computerized Bulletin Board System (CBBS) developed by Ward Christensen and Randy Suess. CBBS officially went online in February 1978 and served as the precursor to what we now refer to as an online forum (Driscoll, 2016; Weyhrich, 2013). The innovators patterned the user interface after the cork bulletin board hanging in their Chicago Area Computer Hobbyist’s Exchange (CACHE) club for posting notices and information to one another. Community bulletin boards are still sometimes seen in libraries, supermarkets, restaurants, schools, and churches.

This same push-pin bulletin board experience for mass communication holds true today with the term “post”now part of the vernacular of online communication. Even though the delivery system has changed, people are people. They still go to the board to see what’s been posted, to discover information that interests them, and to contribute information they believe is valuable to the community.

BBS users fostered community-building and maximized efficiency. Motivation for users was a desire to avoid long-distance call fees incurred by modem dial-up. BBS system administrators used the geographic proximity of users to reinforce the sense of community by hosting local get-togethers. “Online disagreements—flame wars—could be kept in check… because the cost of being a jerk escalated with the likelihood of later seeing your interlocutor face to face” (Driscoll, 2016, para. 15). This same foundation of non-anonymity for PubMed Commons users is an important aspect of participation and a possible factor that affected the forum’s use.

Inevitably, BBS communities developed idiosyncratic personalities and interests, and commenters desired to contribute to an extensive, on-topic conversation with people from a broader reach of society. In 1984, Tom Jennings grew his Fido BBS into a massive 20,000-node network reaching users in South Africa and New Zealand. Within a decade, the FidoNet user base extended to locations throughout the globe. Estimates were that “59% of the nodes were located in North America, 30% in Europe, 4% in Australia and New Zealand, with the remaining 7% split among Asia, Latin America, and Africa” (Driscoll, 2016, para 26).

Acting as centralized locations for topical exchange of information, the BBS communities became opportunities for education and social support. In remote regions of Africa, a FidoNet gateway provided an important means for poorly funded academics to keep current with the latest research. For communities in crisis, BBS like those used by the AIDS Education General Information System (AEGIS) organization disseminated up-to-date research about treatment and prevention that had been found in medical databases (Driscoll, 2016).

One particular BBS feature that contributed to their popularity was that the system served as a file repository. To this day, one advantage of a forum over social media is the ability to archive content so that users can find desired information months or years later. Notably, social media platforms like Facebook continue to develop their “findability” features. The NLM’s commitment to provide a permalink to PubMed Commons comments offered this same added benefit to PubMed users and elevated the status of comments to citable and, perhaps, more formal scholarly communication. As mentioned before, this permanency no longer exists.

As technologies evolve their continued development and utilization is characterized by the way users define and redefine a technology relative to patterns of use, information exchange, functionality, reach, and inclusiveness. In other words, technologies are constantly changing and simultaneously developing as the ICT is iteratively implemented and used (Contractor & Eisenberg, 1990; Haythornthwaite, 2002; Mumford, 1934). Remarkably, the online forum discussion board has demonstrated remarkable resilience in design and functionality since its inception. Online forums are still based on an initial post and the threaded comments offered in reply.

2.2.2 Ongoing Research of Online Environments

Since the first email was sent in 1971, several revolutionary ICTs have developed—online forums being just one. After the World Wide Web was introduced in 1989, there was a surge in research with a goal of better understanding the nature of CMC and how its use could be optimized (Herring, 2004). These pioneer researchers borrowed methodologies from several disciplines to investigate online environments and to increase understanding about community dynamics and their effects on people, organizations, and culture (Preece & Maloney-Krichmar, 2005).

The early research agenda for online forums included ethnographic studies (Baym, 1998; Hine, 2000) designed to discover motivation for use, self-regulation, communication, and lurking behavior; social network analysis for investigating the social structure of an online forum (Wellman & Gulia, 1999; Wosak and Faraj, 2005); content analysis to explore user perspectives (Bauer, 2000; Herring, 2004) and the nature of eLearning posts (Jonassen & Remidez, 2005); online interviews and questionnaires to gain user perspectives (Andrews, Nonnecke, & Preece, 2003); and descriptive studies for reporting recorded frequencies and interactions (Ballantine & Martin, 2005; Romiszowski & Mason, 2004).

Researchers have also investigated content sentiment (Li, Huang, & Zhu, 2010); self-disclosure (Barak & Gluck-Ofri, 2007); interpersonal relationships (Ransom, La Gurdia, Woody, & Boyd, 2010; Steuber & Solomon, 2008); technology adoption (Proctor et al., 2010); and crowdsourcing civility (Lampe, Zube, Lee, Park, & Johnston, 2014). Theories have borrowed from sociology, social psychology, anthropology, and linguistics—their application often based on the disciplinary training of the researchers applying them (Preece & Maloney-Krichmar, 2005).

Notably, research about online information behavior and communication in the context of scholarly communication among biomedical and life sciences professionals is limited in three ways. First, while there are a significant number of studies about CMC hosted on social media and instant messaging platforms (e.g., Sun, Lin, We, Zhou, & Lou, 2018), research about online forums has not received as much attention. Second, the published literature about online forums has centered primarily on social support for health conditions (Casilli, Rouchier, & Tubaro, 2014; Falisi, Wiseman, Gaysynsky, Scheidler, Ramin & Chou, 2017; Glickman, Galhenage, McNair, Barber, Patel, Schulman, & McHutchison, 2012; Hildebrand, Ahumnada, &

Watson, 2013; Seo, 2006); engaging publics with news reports (Montes & Butlier, 2008); product marketing (Kaiser & Bodendorf, 2012; Prendergast, Ko, & Yin, 2010; Xun & Reynolds, 2010); and eLearning (Cheung, Hew, & Ng, 2008; Kahn, 2009). In fact, a great deal of empirical research has investigated online and blended approaches to education and professional development (Smith, 2015). Finally, investigations of online forums utilizing social network analysis to describe relationships among members or topics of discussion are gaining traction as more sophisticated tools for analyses are developed (Kimmerle, Thiel, Gerbing, Bientzle, Halatchliyski, & Cress, 2012; Manca, Delfino, & Mazzoni, 2009; Tirado-Morueta, Maraver-Lopez, & Hernando-Gomez, 2017).

To date, there is a lack of consistent methodology for investigating online forums. Most results are descriptive rather than theory-driven; studies are sparked by excitement over technological novelty (Hilty & Hercheui, 2010). This dissertation aims to fill this gap.

2.2.3 Online Forums as Communities

"Community has become the 'in-term' for almost any group of people who use Internet technologies to communicate with each other" (Preece & Maloney-Krichmar, 2005, para. 1). This precedent was established early on by Howard Rheingold (1993) and Roxanne Hiltz (1985)—pioneers of online community development and research, who used the term "community" to describe the strong sense of fellowship, support, and empathy they observed among individuals participating in the online spaces they studied (Preece & Maloney-Krichmar, 2005, para. 3).

At its core, an online community consists of a) people interacting socially and sharing a purpose, b) policies to guide interactions, and c) computer systems to facilitate a sense of togetherness. According to Malinen (2015), “a community exists in the minds of its members and is constructed symbolically through shared meanings, norms and culture” (p. 229). Despite this widely accepted notion of togetherness, there has been some question through the years as to whether communities—as they have previously been perceived—can actually exist online (Malinen, 2015, p. 229). Not all online forums can be considered online communities, nor have they been designed to coalesce as a community (Blanchard & Markus, 2002). It is important to recognize the wide variety of community types defined by structure, purpose, and users. Many times, any online group is considered a generic, one-size-fits-all brand of community (Gallagher & Savage, 2013). And even those that seem similar can indeed be very different. It is likely that an online forum of peers for the purpose of scholarly communication has different implications than one designed to capture organizational knowledge, rewarding users with enhanced social capital (e.g., prestige, esteem, cachet) for participation (Wosak & Faraj, 2005).

The concept of community has both intrigued and puzzled sociologists, social psychologists, and anthropologists for over a half a century (Wellman, 1982), perhaps even more so as online communities evolve. Respected scholars from varied disciplines have attempted to conceptualize online groups to aid in our understanding. Most have been based on Lave and Wenger’s (1991) idea of a Community of Practice: “a learning partnership among people who find it useful to learn from and with each other about a particular domain” (Wenger, Trayner, & de Laat, 2011, p. 9). Rogers’ (1995) definition of a “social system” would also be appropriate for discussing these cohesive groups: “a set of interrelated units that are engaged in joint problem-solving to accomplish a common goal”(p. 23). Table 2.2 compares concepts of various communities.

Table 2.2

Concepts of Online Communities

Term	Concept	Attributes	Introduced by
Network of Practice	Learning is motivated by demands; Learning is social; and Learning forms identities	A CoP that does not meet face-to-face; show little reciprocity; members have common knowledge and use the Internet communication tools to connect and share information	Brown & Duguid (2000)
Electronic Network of Practice	Individuals are engaged in a human endeavor facilitated through computer mediated communication	A CoP composed of a typically larger, loose-knit, geographically distributed group of individuals	Wosak and Faraj (2005)
Virtual Community of Practice (also "online community of practice"	A shared domain of interest among practitioners or experts participating in a process of collective learning	A CoP developed on and maintained using the Internet	Bara (2010)
Latent Community	"A set of interconnected people who share a common interest on a particular subject even when they have not explicitly disclosed their intent of participation or affiliation" (p. 215)	No formal membership constraints; posts are relevant to a particular subject with members showing an interest in the subject	Yoon, Shin, Kim, Par, & Lee (2009)
Online social gatherings	Online communities that focus on facilitating knowledge sharing among their members		Armstrong & Hagel (1996)

The phenomenon of online communication includes a targeted group of potential users, participating members who post comments, and members who read content but do not participate in the discussion (i.e., lurkers). In online communities, "individual members may not meet or know each other in real life, but they still can share a great deal of what they know with each other" (Brown & Duguid, 2000, as cited in Zhang, 2007).

While online forums have prompted researchers to consider the strength and nature of relationships between individuals to be more useful indicators of cohesiveness than their physical proximity, they've also lured researchers to expend a significant amount of energy trying to define and then characterize which online communities are actually a community (Preece & Maloney-Krichmar, 2005, para. 2). Bruckman (2006) encourages researchers to accept the concept of an online community as one with fuzzy boundaries that is best defined by its membership (e.g., comparing similarities and differences among new and established members of the community). In this way, researchers can focus on the important issues of the organizational processes and the information/communication behavior of the online forum, as well as how the forum coalesces, evolves, or ceases to exist.

An example of this is determining factors affecting a forum's use. Low participation is the most frequently cited reason for the failure of online forums (Ling et al, 2005). Although online forums can support community networks of weakly connected individuals, "the technical implementation needs to be matched with a social implementation to effect connection among yet unconnected others, and to gain a critical mass of communications and users so that connectivity is perceived to exist" (Haythornthwaite, 2002, p. 393). The question to be answered by this dissertation was the extent to which PubMed Commons adequately established the necessary environment for a user community of biomedical and life sciences researchers to support an online forum aimed at promoting scientific discourse.

2.2.4 Biomedical and Life Sciences Researchers as an Online Community

Kuhn (1965) described scientific researchers as a community of individuals engaged in scientific activities. He posited that scientific communities have "undergone similar educations and professional initiations," having "absorbed the same technical literature" with "a subject matter of its own" (p. 177). On the contrary, PubMed Commons represented a diverse mix of many disciplines and stakeholders with different backgrounds. While the PubMed website provided endless opportunities to conjoin an unfathomable number of latent ties (i.e., a connection that is technically available, yet not activated by social interaction), there was a question whether the overall diversity of potential users would negatively affect adoption. This was exacerbated by the fact that possible commenters were limited by the guidelines established by PubMed Commons organizers, including limiting who could contribute, insisting on commenter identification, and moderating comment content.

This discussion makes it interesting to consider the extent to which PubMed users might have eventually developed into a single community, a community of communities, or a *community of independents*. The forum's short life deprived us of this knowledge. Results from a social network analysis (SNA) performed on connections made between 2013 and 2018 are provided in Chapter 4. This type of forum analysis can characterize a network and show the formation of commenter clusters or cliques (i.e., communities). One perplexing question to consider is whether having a shared vocation of research provides sufficient common ground for mutual understanding about a field or motivates commenters to interact with one another. Do disciplinary differences hinder connections? While PubMed Commons members constitute a community concerned with the scientific method of inquiry, there are likely significant disparities in investigative tools, scientific procedures, social norms, and scholarly communication customs that would fragment the group.

This dissertation suggests that although the PubMed Commons online forum was a gathering place in which communities could possibly emerge in the context of a shared interest in scientific research, they simply didn't. If characteristics of the network had been clear, one could assume that forum organizers would have mitigated noticeable holes and structural deficiencies to increase the rate of adoption and use. Research has shown that relationships are critical for obtaining information (e.g., Burt, 1992; Granovetter, 1973; Rogers, 1995; Szulanski, 2000) and learning how to do your work (e.g., Brown & Duguid, 1991 & 2000; Lave & Wenger, 1991; Orr, 1996; Wenger, 1998). Informal communication outside institutional constraints and controls is known to be a catalyst for innovation and a means for disseminating new knowledge (Cothrel & Williams, 1999).

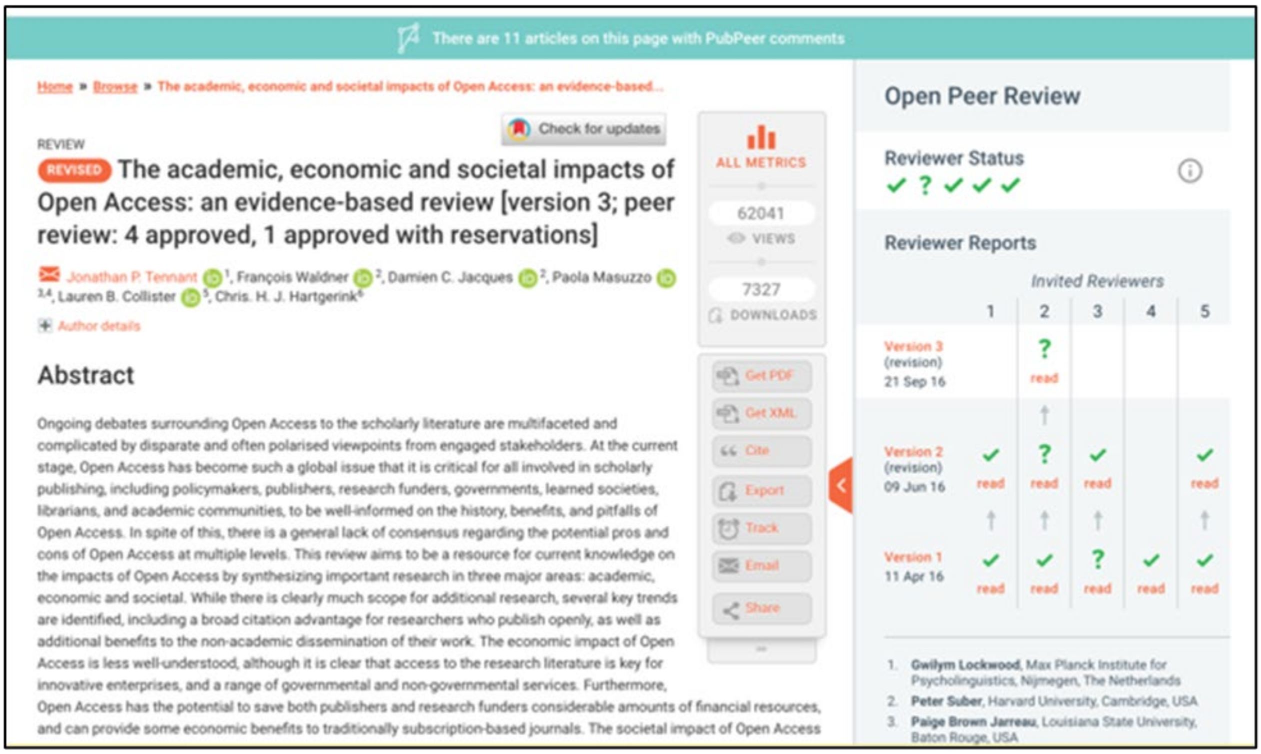
2.3 Scholarly Communication

Scientific and scholarly communication has captivated scholars for centuries (Meadows, 1998). Current developments are no exception. In fact, scientific communication is at a critical juncture of transition—a crisis in publishing mixed with a weakness in the perceived value of the peer review system (Proctor et al., 2010). Furthermore, there is an impending sense that the dissatisfaction with traditional publishing and the possibilities introduced by new technologies (e.g., nearly instant transfer of information) can activate a more open mechanism for scientific communication (i.e., open science) (Fjällbrant, 1997).

Nearly 50 years ago, William Garvey and Belver Griffith characterized a system of scientific communication among a community of psychologists by its informal and formal communication behaviors. Garvey and Griffith (1972) proposed that communication is the essence of science, and that scientific communication is a social process. Meadows (1998) affirmed this, saying, Communication is at the heart of research. It is as vital for research as the actual investigation itself, for research cannot properly claim that name until it has been scrutinized and accepted by colleagues. This necessarily requires that it be communicated. (p. ix)

Figure 2.1

Digital Format Indications of Peer Review and Post-publication Metrics



Screenshot of an article on the F1000 website, which shows evidence of peer review versions, post-publication metrics, and an API message that comments are available on PubPeer.

The Garvey-Griffith model of the scientific communication system implied that the refereed scientific article was the product of the process of communicating research in various stages ranging from the initial concept to the integration of the research as an accepted component of scientific knowledge (Hurd, 2000). The advent of Internet-based ICTs has radically increased options for communicating that will likely lead to several new ICTs aimed at scientists. Teasing out which ones take hold will take time. History has repeatedly shown that technical possibilities are not always embraced by the majority (e.g., microfilm for replacing paper in libraries, beta videocassette recorders for providing an alternative to watch movies). Inconveniences and group norms have repeatedly fueled resistance to innovation. These concerns are echoed in sentiment about Internet-based scientific discussions and PPPR. Hurd (2000) predicted that “peer review will be a feature of any new communication system, although the mechanisms to ensure quality may differ in a digital submission and review process” (p. 1281). As technologies have advanced, different forms of peer review continue to evolve. Consider Figure 2.1, which shows an article that has undergone a variety of peer reviews and alternative metrics to report its effectiveness.

An exponential increase in scientific inquiry, scholarly publishing, and university centers introduced an increasingly robust research agenda in the 1960s. The result was termed “scholarly communication” in the mid-1970s. “Researchers sought to understand the processes involved in scholarly communication by building models of information flow and by testing theories of behavior” (Borgman, 2000, p. 412). A seminal book by Meadows (1974) examined “how and why scholars do research; how they communicate with each other; how, when, why, and where they publish; and how publishing and libraries interact with scholarly practices” (Borgman, 2000, p. 413). Interest in this strain of research saw a resurgence in the 1990s due to the emergence of ICTs. Since then, scholarly communication has come to be conceptualized as:

> the system throughout which research and other scholarly writings are created, evaluated for quality, disseminated to the scholarly community, and preserved for future use. The system includes both formal means of communication, such as publications in peer-reviewed journals, and informal channels, such as electronic listservs. (Association of College & Research Libraries, 2003, para. 1)

Borgman (1990) emphasized the social processes, saying that scholarly communication "includes the growth of scholarly information, the relationships among research areas and disciplines, the information needs and uses of individual user groups, and the relationships among formal and informal methods of communication" (p. 14). Proctor et al. (2010) further explained that scholarly communication involves "communicating scholarly ideas to broader [scientific] communities" (p. 4040), each with its own practices and cultures. "These disciplinary and local cultures have a strong influence on how new information and communication technologies (ICTs) are adopted" (p. 4040).

Bubela and Caufield (2009) characterized scientific communication as "a complex and contentious topic that encompasses a spectrum of issues from the factual dissemination of scientific research to new models of public engagement" (p. 514). These are exacerbated in the 21st century with changes in technological, social, and institutional policies, as well as challenges to traditional publishing models. New media are fundamentally changing the nature of scientific communication, and with that, discussions are needed about new modes of online digital formats that blend research reporting with user-generated content (Bubela & Caufield, 2009).

This urgency is even more critical given events that have propelled a dramatic increase in academic output (e.g., increased specialization, new technologies and methods, an interdisciplinary approach to scientific discovery, and the institutional mindset for publish or perish). "In the last two centuries, the number of scientific articles has doubled every 10-15 years" (Bravo, 2014, p. 5).

2.3.1 Formal and Informal Scholarly Communication

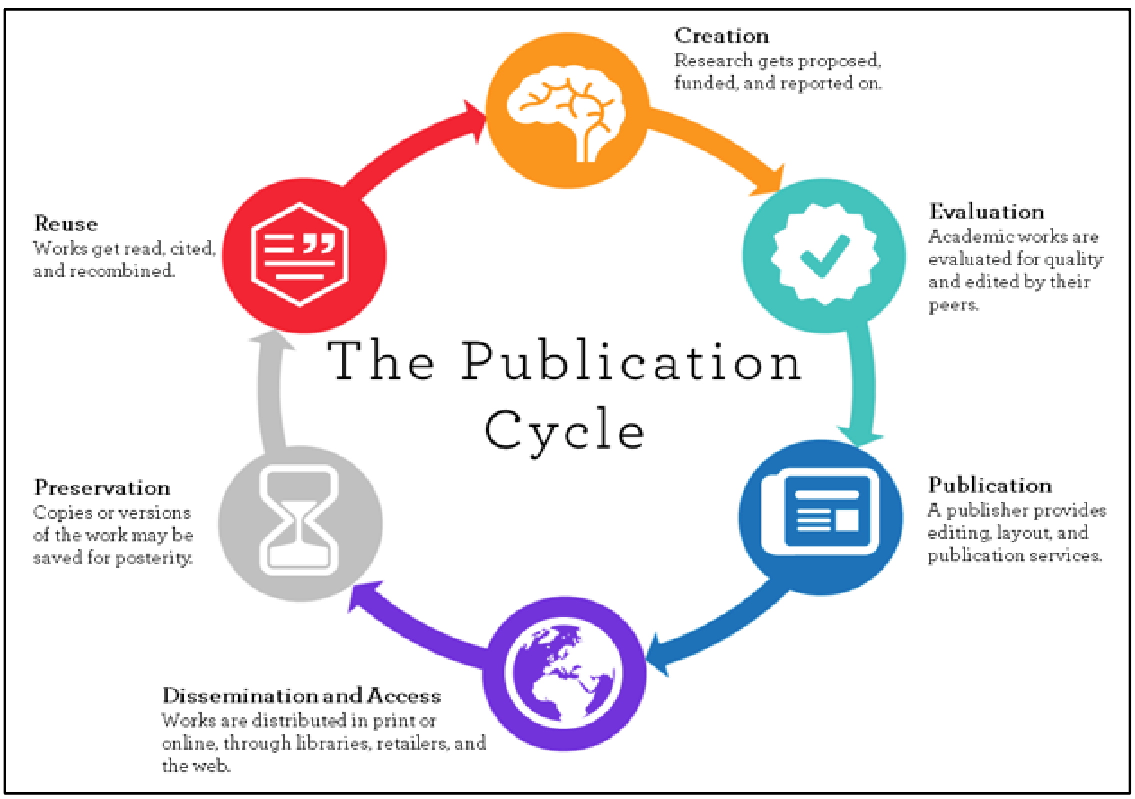
Although the line between formal and informal scholarly communication is beginning to blur, a distinction between the two is necessary for this dissertation. In general, scholarly output fed by informal communication progresses to developing formal communication outputs, such as conference proceedings, journal articles, and then abstracts. Indexed metadata, edited volumes, books, and reference sources follow soon after (Bubber, 2013).

2.3.1.1 Formal Scholarly Communication

Peer-reviewed publications have been the hallmark of formal scholarly communication. These products are the output of the formal publication cycle, which includes a spectrum of scholarly activities ranging from creation and evaluation to publication and reuse (see Figure 2.2).

Figure 2.2

The Publication Cycle



The Washington University Library's (n.d.) visualization of the publishing lifecycle.

Evaluation has historically occurred through peer review—the process by which publications and presentations are evaluated by others who are practicing members of a particular field. Letters to the editor, expressions of concern, errata, and retractions serve as formal methods for peer review after publication (i.e., PPPR). Although the scientific journal article has been widely favored to diffuse scientific and technological information for nearly 500 years, the Internet and ICTs are changing this.

2.3.1.2 Informal Scholarly Communication

Informal communication is difficult to measure. Historically, the communication of observations and new experiments were exchanged verbally at the meetings of learned societies and in personal letters between individual scientists and research groups. "These were sent to a person who acted as a 'gatekeeper' for transmitting news" (Fjällbrant, 1997).

Building on the work of de Solla Price (1963), Crane (1972) used the term "invisible college" to describe researchers engaged in these types of informal social exchanges. A true visionary, de Solla Price suggested more than 50 years ago that using journals for scientific communication among colleagues is ineffective and outdated. From his perspective, the sheer speed and growing volume of scientific information had caused books to give way to journal papers, and then, in turn, to letters to the editor. Price intimated that mid-20th century scientists at the forefront of research did not read, but rather telephoned, discussed at society meetings and conferences, and developed invisible colleges of small groups of peers to share information. He had a notion that the traditional paper trail of scientific reporting was little more than tallying up the number of papers a scientist was required to publish. He held up Barnaby Rich (c. 1613) as an example of a scholar who understood the original purpose of scholarly writing, claiming it was "a social one of finding out what was being done and by whom rather than a scholarly one of publishing new knowledge" (de Solla Price, 1966, p. 63).

Irrespective of this foretelling, the practice of scholarly bricklaying via publishing and citations remains a tradition in scholarly work and communication, even more so as protection of intellectual property intensifies (de Solla Price, 1966, p. 65). According to de Solla Price, paper writing became more about staking territory—monitoring rivals and peers—and less about communicating new knowledge. He suggested that as Little Science gives way to Big Science, scientists have been and will be urged to write, not read.

Current modes of informal scholarly communication include some that de Solla Price described (e.g., conference meetings and phone calls); however, due to ICTs, these also include emails, blogs, social media alternatives, and online forums.

> Computer mediated communication is providing a technological basis for new forms of spatially dispersed, loosely bounded, networks of scholars that are more connected than the fitful, amorphous relationships of the past and less physically proximate and bureaucratically structured than contemporary universities. The velocity of communication is more rapid, distant scholars stay in touch more, and email and attachments fill gaps between face-to-face meetings. (Koku & Wellman, 2002, p. 3)

2.3.2 Peer Review

The following discussion about the peer review process is necessary, as the purpose of PubMed Commons was often misconstrued as strictly a vehicle for PPPR. Chapter 5 provides an extensive history and explains that forum innovators like Tibshirani envisioned a place where scientists could exchange ideas, ask questions about methods/techniques, offer suggestions, make comments on each other's work, and even offer praise (Spector, 2013, para. 7).

Current unrest in academic publishing has people questioning the processes for peer review. Traditionally peer review has been considered the "gold standard by which academic manuscripts are vetted for publication" (Herron, 2011, p. 2275); however, the how and when of this task are changing. Peer review can be conceptualized as the process by which formal scholarly communication is evaluated by others who are practicing members of a particular field. It's important to acknowledge that historically, this occurred prior to publication. Thus, the foundation of peer review was based on informal scholarly communication, which was affected by an increasing size of scientific community.

> In the early days of scientific societies (i.e., the 17th century), scientists would share their experimental results with each other at meetings and receive feedback about their experiments in person. As the scientific community grew, it was impossible for everyone to be in the same room to hear about results, and so the amount of immediate feedback offered was limited to a few conferences or other gatherings. Recently, publishers, scientific societies and entrepreneurs have begun using the web to bring back the era of immediate feedback: so-called 'post-publication peer review'. (Swoger, 2014, para. 1)

This gives credence to a growing interest and participation in pre-publication peer review opportunities (e.g., arXiv, bioarXiv, PrePubMed). Recall that the NLM introduced a system for pre-publication in 1960, disrupting the process of peer review at the time.

Many scholars will agree that a process of peer review benefits all stakeholders—publishers, editors, authors, and reviewers. Editors rely on the feedback from reviewers to inform choices about competing manuscripts submitted for publication. Authors gain insights about ways to improve their research output and improve the quality and clarity of their manuscript. Reviewers are rewarded with developing expertise in their field and being recognized for their contribution to the profession.

In its intended form, peer review is espoused to be honest and beneficial for advancing scientific discovery. It can be either a closed or open process. Closed is accomplished by single-blind or double-blind review. The former ensures that the author is not aware of the reviewer's identity, although the reviewers are aware of the author's identity, affiliation, and credentials. The latter ensures that neither the author nor reviewers are aware of each other's identity. Single-blind review has been criticized for possible reviewer bias, unethical infractions against intellectual property, purposeful delay, and misleading feedback. Double-blind review eliminates chances of bias, although experts in a field are likely to recognize authors' work.

In open peer review, authors and reviewers are clearly known to each other. Their names might even be published alongside each other, with reviewers' reports printed in the final manuscript. Proponents claim that removing secrecy ensures intellectual property rights and discourages reviewers from offering careless or offensive comments. Instead, they are recognized for their contribution to the process of advancing science. Opponents believe that full disclosure might in fact insert bias and encourage overly critical and inappropriate comments.

PubMed Commons guidelines were designed to reduce and even mitigate negative experiences of other online forums, like PubPeer, which permit anonymous comments. In fact, PubPeer has been named in a defamation court case (Science News Staff, 2016). I claim that PubMed Commons guidelines for open review by any author/researcher, regardless of their expertise, on the work of another author's work was indeed a factor that affected the forum's adoption and use. To the credit of the PubMed Commons team, the intent was to keep the comments as seemingly unbiased as possible while also protecting commenters from litigation (H. Bastian, personal communication, October 28, 2016).

In 2015, Bosman and Kramer reported 101 Innovations in Scholarly Communication, including discovery, analysis, writing, publication, outreach, and assessment. Of publications, they noted more use of ‘publish first, judge later,’ and of assessments, they shared an expectation for more open PPPR.

Given technological advances, PPPR can occur rapidly. Online scientific venues similar to PubMed Commons (e.g., F1000, ResearchGate, Publons, PubPeer, RetractionWatch), publisher specific journals (e.g., *PLoS One, British Medical Journal, Nature, New England Journal of Medicine*), personal blogs, and even social media sites (e.g., Twitter) facilitate PPPR to some degree in spite of inherent issues that threaten their widespread acceptance, namely anonymity, fragmentation, and qualification of commenters.

Advocates of PPPR are quick to point out its benefits. The process is more transparent, and it may promote more rigorous, tactful, and constructive comments, as reviewer names are openly known. Also, a wider group of people are empowered to comment on the paper.

Forums like PubMed Commons and similar Interactive web technologies like those mentioned above are indicative of the foreseeable changes in scholarly communication. Certainly, PubMed Commons editor Bastian (2014) was an advocate for a culture that accepted a more open science, which would require a change in the traditional publishing cycle: “both improving research quality and reducing waste in science require a stronger post-publication culture” (p. 1). Chapter 5 discusses how Bastian’s role in this movement affected PubMed Commons.

2.4 Chapter 2 Summary

A meaningful investigation of organizational processes and the information/ communication behaviors that characterized the PubMed Commons online forum requires a working knowledge of the many characteristics and influences that were inserted by its formation, stakeholders, and current societal norms. This chapter provided necessary background information about the context of PubMed, the history of online forums, and current issues in publishing that affected PubMed Commons use.

CHAPTER 3

METHODS

The retrospective explanatory case study research of PubMed Commons was designed to discover how the forum functioned as an online commenting system for scholarly communication and what factors affected its use and disuse. Data analysis was intended to shed light on the organizational processes for developing the online forum and to explicate the information behavior and subject matter of naturally occurring communication among scientists captured between June 12, 2013, and Feb. 15, 2018. Guided by an interpretivist approach to inquiry, this dissertation used mixed methods to ensure scientific rigor.

Social scientists are motivated by their desire to discover more about all aspects of society, especially social systems, relationships, and individuals' behaviors. They aspire to produce replicable results and illuminate social meaning of the world rather than factual aspects of it. Information studies focus on the way people access, store, retrieve, and use recordable information, as well as the technologies and related services that facilitate the management and use of the information (Association for Information Science and Technology [ASIS&T], 2021).

Researchers are guided by their philosophical worldview, which informs their rational attempt to explain life. A post-positivist outlook is deterministic, operating from a belief that cause and effect are the pathway to unveil absolute "T"ruth. Ontological foundations of this worldview suggest that social reality is separate from the researcher; its epistemology promotes a single social reality, even though this is sometimes difficult to access. Validity is achieved by consistency of measurement, and results are a value-free (i.e., neutral) assessment that can be generalized and replicated. Quantitative measures and analysis provide a more objective reporting of regularities.

By contrast, an interpretivist reality is socially constructed and informed by people's lived experiences and shared meanings. This philosophical worldview supposes "t"ruth is intersubjective and constantly changing (Miller, 2000). Ontological foundations advocate that together, study participants and researchers construct meanings of realities embedded in experiences. The researcher and his or her perceptions serve as study mediator (i.e., the tool for investigation). Epistemologically, interpretivists believe social reality is multifaceted and assuredly fractured. The goal of inquiry is not a-contextual insight; rather, claims of contextualized knowledge resonate with lived experience and can be stated at a level of generality so findings are transferable in other contexts. Through credible research, the researcher seeks probability rather than certainty. Furthermore, no single or correct route can lead to truth. Knowledge and meaning are found through in-depth observations using various techniques. Factors like context, individual choice for decision-making purposes, and intentions are active agents. Dependability is vital and is achieved through a trustworthy researcher who knows participants' experiences. The objective is discovering sensitizing concepts and social dynamics through meaning-making, framing, and throughputs.

Guba and Lincoln (1994) explained that an inquiry paradigm is constructed by methods for investigation that are coupled with philosophical worldview. Tools for discovery merely provide strategies for researchers to examine a phenomenon. Both quantitative in the post-positivist tradition and qualitative in the interpretivist tradition afford a scientifically rigorous way of accomplishing inquiry based on empirical (i.e., observable) reality. Study design is what renders results largely transferable/applicable in other contexts.

This chapter details how the investigation for this study was conducted, including its purpose, participants, and methods for data collection and analysis. Given the abundance of online forums, there is surprisingly minimal empirical research into their organizational processes and the assumption that information shared among participants achieves sponsors' stated purpose or satisfies user needs. The lack of direction about how to develop and evaluate a forum is in sharp contrast to anecdotal information about them. While large data sets (i.e., big data), like comments in online forums, can be analyzed computationally to uncover associations, patterns, and trends in behavior, these methods offer little help in contextualizing the phenomenon or revealing the multifarious factors that affect its purpose, adoption, and use.

As participation in online forums began to grow in popularity, online information researchers Eysenbach, Powell, Englesakis, Rizo, & Stern forecast an ongoing need to evaluate under which conditions and for whom online forums are effective and how effectiveness can be maximized (2004). Studies in online forums used in education and health care social support are abundant; however, investigations directly aimed at developing systems for scientists to engage in scholarly communication are just starting. More should be known about the ways in which ICTs are implemented and used for this purpose. This retrospective explanatory case study research about PubMed Commons provides an excellent contribution to this understanding.

3.1 Purpose for Researching PubMed Commons

The National Institutes of Health (NIH) Public Access Policy executes Division F, Section

217 of the Omnibus Appropriations Act of 2009, ensuring that the research community and public at large has access to results and activities of NIH-funded research (NIH Office of Extramural Research, 2017). The National Science Foundation (NSF) Public Access policy was enacted in 2013 and requires organizations receiving awards on research proposals submitted or due on/after Jan. 15, 2016, to make publications available in the NSF Public Access Repository (PAR) (NSF, n.d.). Public access policies ensure that U.S. government-sponsored research is digitally archived in full-text format. Articles receiving NIH support are available on the National Library of Medicine's (NLM) PubMed Central database with corresponding abstract and citation entries indexed on the free PubMed search engine database.

Understanding the gravity of this archival responsibility and the direct impact biomedical literature has on health care for humans and animals, NCBI director Lipman joined an enthusiastic group of scientists to conceptualize and introduce PubMed Commons in 2013 (Spector, 2013).

As a research specialist working with medical students and faculty at the Oklahoma State University Center for Health Sciences, I became increasingly interested in PubMed Commons comments that began to appear in PubMed. Accordingly, this dissertation was designed to investigate the organizational processes of how PubMed Commons was developed, as well as the information sharing/communication behaviors that were naturally occurring in the comments posted below the abstracts on the PubMed GUI. The goal of this dissertation was adding depth to the existing body of knowledge and humanistic understanding of how individuals utilize Internet-based, ICTs for scholarly communication in an online environment. Thoughtfully designed research questions, appropriate theoretical sampling, and inductive reasoning were aimed at reducing the gap between the research and its contribution to social-scientific theory and practice. Because the use of online forums for scholarly communication is a phenomenon not well explained by theory, rigorous inference-making was made in Chapter 5.

Data for this dissertation was carefully collected from a variety of sources, which followed precedence for systematic processes and procedures required of a case study. This research method of inquiry is well-established in the social sciences (Yin, 2009). Longitudinal data collection and analysis permitted necessary overlap. Results from a pilot study of PubMed Commons I conducted with Burns during graduate school in 2014-15 served as a foundation for my research and was used to inform decisions about research questions and methodology for this dissertation.

3.2 Methodology for Answering Research Questions

A research question is simply the question that motivates a researcher to embark on their journey of discovery. It guides all stages of inquiry, data collection, analysis, theory building, and reporting. The following research questions framed this retrospective explanatory case study research.

RQ 1: How did PubMed Commons function as an online forum for posting comments about published articles?

a. How often and in what ways did stakeholders participate in PubMed Commons?
b. What were characteristics of forum commenters on PubMed Commons?
c. What was the subject matter of comments posted on PubMed Commons?
d. What types of communities formed on PubMed Commons?

RQ 2: What factors contributed to or limited scientific discourse on PubMed Commons?

RQ 1a was answered through a quantitative content analysis of comments posted on article abstracts indexed in PubMed. The collective whole of comments posted between June 12, 2013, through February 3, 2018, constituted “PubMed Commons.” RQ 1b was answered by creating a database to find associations between commenter behavior on the forum and commenter characteristics at the time the forum closed. RQ 1c was answered by performing a qualitative content analysis of comments posted on PubMed Commons. RQ 1b and RQ 1c used a statistically significant random sampling of commenters and their comments. Details are provided below. RQ 1d was answered by performing a social network analysis (SNA) of all relationships that formed between commenters on PubMed Commons during the 4½ years the forum was operational.

RQ 2 was answered following RQ 1 data collection and analysis. Results were informed by a key informant interview with PubMed Commons editors Hilda Bastian and Melissa Vaught and supplemented by a systematized search for articles, gray literature, and social media communication about PubMed Commons.

3.3 Participants

The scope of this study included information about the behavior of 1,551 published scientists whose articles were indexed in PubMed and who made comments between June 12, 2013, and February 15, 2018 on their own or other scientists’ publications. Note that although the official last day of the forum was Feb. 3, 2018, comments were accepted and visible through Feb. 15, 2018. Sample size for qualitative content analysis was n=381.

3.4 Research Design

Creswell (2007) reminds scholars that rigorous research and its validation is a process.

Reliability and validity for this study were achieved using the following advice from Creswell: a) be impeccable about procedures (e.g., detailed records, reliable software); b) employ deliberate strategies for triangulation using a number of different sources, methods, investigators, and theories; c) seek external validation through informal peer review and study participant reflections; d) perform negative case analysis to identify outliers/exceptions and to challenge the researcher's initial hypothesis; and e) ensure findings have resonance that readers can relate to (pp. 207-209).

Interactive web-based technologies have added new opportunities and challenges for social science research. A mixed method approach was necessary to capture the interdependence of people, processes, and technology at play when examining PubMed Commons and its organizers/commenters. Iterative data collection and analysis of quantitative and qualitative data was used to exploit advantages and mitigate risks associated with using any one data type or method. Triangulation facilitated validation through cross verification of data from multiple sources and ensured rigorous and comprehensive research about PubMed Commons. Four widely accepted research methods for information scientists were used. These included: 1) quantitative content analysis, 2) qualitative content analysis, 3) semi-structured interview, and 4) social network analysis. Combined, the quantitative and qualitative analysis of the forum were considered "computer mediated discourse analysis (CMDA)."

Data came from a variety of sources. A .csv file of archived comments and commenters who posted on PubMed Commons between June 12, 2013, and Feb. 15, 2018, was downloaded from an NCBI FTP website. Historical PubMed Commons data were also captured from the PubMed GUI on Feb. 18, 2018, following a PubMed search for all records (i.e., "all[sb]") with

PubMed Commons comments. Results were downloaded; screenshots of the forum GUI were taken; and comments with associated information (i.e., threaded/nested comments, helpful/not helpful ratings) were copied from the PubMed GUI and pasted onto MS Word documents. A key informant interview with PubMed Commons editors Bastian and Vaught was conducted Oct. 28, 2016, at their offices in Bethesda, MD. A social network analysis of the relationships that formed between all commenters on PubMed Commons between June 12, 2013, and Feb. 15, 2018, showed the extent to which relationships formed within the forum. Finally, a systematized search and review of published articles, gray literature, and social media content that focused on PubMed Commons was iteratively conducted between April 2014 and June 2021. This information provided necessary context for the case study. More information about each of these data sources and the methods used for analyzing them is detailed below.

3.5 Methodologies for Data Collections

"Research approaches are plans and the procedures for research that span the steps from broad assumptions to detailed methods of data collection, analysis, and interpretation" (Creswell, 2007, p. 3). Social scientists utilize a variety of systematic techniques for providing contextual information to provide rich insight into human behavior and uncover an insider view of a studied phenomenon (Guba & Lincoln, 1994, pp. 106-107).

The comprehensive, multifaceted investigation of PubMed Commons was informed by data collection and analysis using combined methods for examining users and associated information behavior (i.e., commenting, responding), as well as the processes undertaken for introducing the forum on the PubMed GUI. This study was approved by the University of North Texas Internal Review Board (IRB).

3.5.1 Computer-Mediated Discourse Analysis

When examining an online forum, the content of what is being shared and how it is communicated are factors in who participates. These cues help possible adopters better understand the purpose of the forum and assess complexities of its use. The importance of capturing communication as it naturally occurs cannot be underestimated. The term "computer-mediated discourse analysis" (CMDA) has been used to explain the scientific approach to studying computer mediated communication (CMC) carried out over computer networks or wireless technologies.

> [CMDA] is often supplemented by surveys, interviews, ethnographic observations, or other methods; it may involve qualitative [rich, thick descriptions of a phenomenon] or quantitative [statistical based on counts] analysis; but what defines CMDA at its core is the analysis of logs of verbal interaction (e.g., characters, words, utterances, messages, exchanges, threads, archives, etc). In the broadest sense, any analysis of online behavior that is grounded in empirical, textual observations is CMDA (Herring, 2004, p. 339).

CMDA aims to measure variables without manipulating independent variables. This method is unique in that much of the social context has been stripped away. Typically, such factors present problems (i.e., controlling independent variables is difficult in complex interactions). Results from my pilot study of PubMed Commons indicated that both quantitative and qualitative methods should be used to gain an exhaustive overview of the forum and to discover insights about commenters and their commenting behavior.

A major challenge to CMDA is determining which elements to measure and how to code them. Quality CMDA results cannot be achieved without careful attention to a researcher's coding practices. Often in quantitative CMDA only the validity of a researcher's guidelines is reported, not their application in the actual process of establishing a coding scheme or initial analysis of the content (Hak & Bernts, 1996, p. 231). In this dissertation, developing explicit coding instructions and appropriate codes for quantitative and qualitative data collection ensured that inferences about specific characteristics of interest (e.g., single posts, threaded responses, nested threads, moderator involvement, and commenting behavior) were captured and reported with confidence. These characteristics are operationalized further in Chapter 4.

A code book for analyzing qualitative comment content was developed from previous studies of PubMed Commons, including the pilot study I conducted with Burns (see Appendix D), a study by PubMed Commons editors Bastian and Vaught (see Appendix G), and a study by Lane (2016a). Categorization (i.e., coding) and analysis for each phase of the qualitative content analysis portion of the CMDA was iterative and aided by grounded theory (Glaser & Strauss, 1967). Codes were exclusive, exhaustive, and equivalent, which ensured that 95% of message thoughts in the final dataset were codable. Message thought units of analysis are described under the qualitative CMDA content analysis section below.

3.5.1.1 Quantitative CMDA Content Analysis

Numeric CMDA reported frequencies and intensity measures of forum characteristics. Data were separated into comment-level, article-level, journal-level, and forum-level groupings. Examples of defining attributes for comment and commenter codes are listed below. See Chapter 4 for complete data coding nomenclature and results.

- Publishing date of article receiving comment
- Number of comments (i.e., single, 2, 3, 4, 5, 6 or more) posted on article
- Number of journal club comments vs. individual commenters
- Number of articles commenter posted on
- Relationships between commenters

- Number of comments removed by moderator or edited/deleted by commenter
- Number of helpful/not helpful ratings
- Number/name of journals with articles receiving comments
- Time of day/day of week comments were made

3.5.1.2 Qualitative CMDA Content Analysis

One shortcoming of quantitative CMDA is that findings are not particularly interesting. Qualitative CMDA extends the quantified observable content by intensely examining texts and categorizing them into verifiable units of analysis with similar meanings.

PubMed Commons comments were sorted on Excel spreadsheets and coded into "message thoughts" by two independent investigators. In 2000, McKenzie and Murphy (as cited in Hew & Cheung, 2003) described message thought units of analysis as discrete ideas relating to a specific topic. Accordingly, text surrounding an operationalized concept in each post was analyzed for its direct relatedness.

A random sampling of PubMed Commons comments captured variability and made findings generalizable to a larger ration of the forum. A sample of 381/7,629 comments were analyzed to achieve a 95% confidence level with 5% margin of error. The sample size ensured a standard deviation z-score of 1.96.

Although CMDA of online texts mitigates some risks to validity associated with self-report on surveys and in interviews, as well as the absence of real-world influence in experiments, it should be noted that the CMDA process of coding reinserts risk. For this reason, CMDA is merely one method I used for investigating PubMed Commons.

3.5.2 Commenter Characteristics

Commenting behavior and commenter demographics were entered as descriptors into a database I created on the Dedoose web application (see Appendix C). Dedoose was developed by academics from UCLA with support from the William T. Grant Foundation. The software as a service (SaaS) platform is designed to help researchers organize, analyze, and find insights in unstructured data (e.g., interviews, open-ended survey questions, published literature, social media posts, and web content). Interested readers can learn more about Dedoose at dedoose.com. Characteristics of commenters were gathered from several information resources. One source was the archived “commons_archive.csv” file, which was accessed via the NCBI FTP website. Another was downloaded records from a PubMed search for “all[sb]” with the “Reader Comments” filter activated. These data were supplemented with information freely available from various searches on the PubMed search engine and on the Internet.

3.5.3 Social Network Analysis

A social network analysis (SNA) of archived data from two information resources used to inform about commenter characteristics (i.e., “commons_archive.csv” file and “all[sb]” PubMed search records) was used to map the social structure of participants on the forum and determine to what extent communities formed on PubMed Commons.

SNA is both an approach (i.e., theory) and a set of techniques (i.e., methodology) that can be applied to study the invisible relational structure of a group (Schultz-Jones, 2009, p. 595). SNA theory aims to explain the workings of networks as generalizations about relationships among variables; SNA methodology is a systematic approach for data collection and analysis. Social structures of a community allow a researcher to generalize about key individuals and community-forming processes. Knowing shared similarities, social relations, interactions, exchanges, and flows among commenters in an online forum network can aid organizers in assessing forum rate of adoption, and then restructuring the network as needed (Wassermann & Faust, 1994). SNA is an important tool for predicting collective action (Burt, 1992) and influencing the willingness for knowledge sharing among the group (Rogers & Kincaid, 1981), as well as promoting innovation, improving efficiency, and contributing to desired changes in an organization's culture (Kline & Sanders, 1993). SNA can be conducted during separate phases of a forum's adoption or to compare one forum's network with another.

The study of social networks is a rapidly expanding, multidisciplinary area of scholarly pursuit involving social, computer, statistical, and mathematical sciences. Information scientists will recognize citation analysis as a form of SNA for examining formal networks imposed by documented connections of citations and co-authorship. SNA has also been used by researchers in this field to study ways for improving access to information, assessing knowledge sharing in informal groups and organizations, and discovering how information is spread in scientific communities (Haythornwaite, 1996; Marion, Garfield, Hargens, Lievrouw, White, and Wilson, 2003; Schultz-Jones, 2009).

All commenters in this study were considered part of a large sociocentric (i.e., whole) network. Forum data was formatted into a one-mode, nodelist adjacency matrix populated with PubMed Commons' commenter names on both the x and y axes. Symbols for directed and valued ties aided in determining clusters, cliques, centrality, and density. Directionality was determined by recording if a commenter responded to a comment. Valued relations (i.e., number of posts each commenter made) demonstrated commenters' roles in the network. A network sociogram (i.e., visualization of algorithmic results) showed social cohesion of clusters and cliques. Lines (i.e., ties) connecting commenters (i.e., nodes) in an unfiltered network sociogram shown in Chapter 4 were weighted to aid in visualizing relationship strength (Baxter, De Reiemer, Landini, Leslie, & Singletary, 1985). SNA measures demonstrate why mixing data collection techniques in this dissertation (e.g., CMDA quantitative and qualitative content analysis with SNA) resulted in a more comprehensive understanding of PubMed Commons.

3.5.4 Key Informant Interview

A key-informant interview and email correspondence with PubMed Commons editors Bastian and Vaught provided understanding and insights about forum development, guidelines, promotion, moderation, and assessment, as well as commenter information/communication behavior and the forum's perceived positioning in scholarly communication among biomedical scientists.

Qualitative interviewing requires listening and hearing (Rubin & Rubin, 2005). Accordingly, I came prepared for the October 28, 2016, meeting with a planned list of questions informed by my 2014-15 pilot study with Burns (see Appendix F). My colleague Julia Crawford accompanied me to the interview and took extensive notes during the meeting. I later transcribed answers into an MS Word document.

3.5.5 Systematized Search for Historical Data

A systemized search and review of articles, formal presentations, gray literature, and social media content about PubMed Commons was conducted every six months, beginning in

April 2014. NCBI Insights and PubMed Commons Blog entries were part of data collection. Information was iteratively reviewed to provide context to data analyses. Examples from the systematized search are provided as evidence throughout the discussion in Chapter 5.

Twitter Web Analytics. A Netlytics-supported network analysis on Nov. 15, 2015, collected baseline data for the PubMed Commons Twitter account. Early network relationships were compared with data collected on Jan. 1, 2018. A sociogram visualized how tweets and retweets about PubMed Commons were diffused throughout the Twitter platform. A word cloud of the top thirty words on each date showed that comments were centered on scholarly communications. This information is presented in Chapter 5 as evidence of PubMed Commons' organizer-driven forum promotion and PubMed Commons' commenter topics of interest.

3.6 Methods for Data Analyses

This section describes the methods I employed when analyzing the collected data described above.

3.6.1 Quantitative Computer-Mediated Discourse Analysis

Quantitative CMDA content data were analyzed through iteratively sorting and manually counting various comment-level, article-level, journal-level, and forum-level characteristics and postings. Numerical and statistical analyses were used to illuminate additional features that should be counted to gain a more complete understanding of the forum. Details about sorting and calculating data are available in Chapters 4 and 5.

Additional quantitative analyses were performed algebraically using Dedoose. Results were based on 25 commenter descriptors that were input for each of the 381 participants included in the sample set. I employed my colleague Crawford and a student assistant to help me enter specified data into the database. I uploaded PubMed Commons comments from an Excel spreadsheet as a separate media file so that each comment could be associated with characteristics about the person who made the comment. Separate reports were run for various descriptors, and the most significant results are depicted in Chapter 5.

3.6.2 Qualitative Computer-Mediated Discourse Analysis

Comment categories for message thoughts were developed using a grounded theory approach. Grounded theory data analysis helps researchers avoid the “everyone knows that” effect of social science research and provides researchers with tools to study continuously changing processes within the social world. Each unit of data (e.g., message thoughts) was placed into as many categories as possible, depending upon classifications of data that emerged and the way in which a message thought fit with an existing category (Glaser & Strauss, p. 105). This method is commonly referred to as constant comparative analysis because units of analysis are constantly compared with each other as the analysis progresses from one stage to another. The process begins with the researcher reading through the data set and eliminating data that is not relevant to the phenomenon under study. The remaining data advances from open coding—when categories are expanded—to focused coding—when categories become more selective and are contracted. The final stage is axial coding at which time inter-related themes can be established. During this stage, negative cases that do not fit in an established category are identified.

The original codebook for PubMed Commons comments was developed for my pilot study by examining 232 comments made between June 2013 and April 2015. After reading through the dataset and removing comments that were not relevant to the study (e.g., comments removed by moderators), I performed open coding and proposed themes to Burns. Coding syntaxes were reviewed and discussed until a saturation of mutually agreed-upon categories was developed. Axial coding prompted further data analysis, causing us to link properties and dimensions of some categories and create subcategories in others. Each code was explicitly defined. Comments in the dataset were coded independently, and conflicts were resolved. Eleven codes emerged from the data and were used to describe PubMed Commons comments. These were reported in a poster presentation at the 2015 National Conference on Health Communication, Marketing and Media in Atlanta, GA (see Appendix D).

The pilot study codes served as the foundation of data coding for this dissertation. First, I compared my pilot study code book with codes I discovered had since been developed by Bastian and Vaught (2016) and Lane (2016a) in their investigation of the forum. I synthesized and regrouped codes, updated definitions, and generated a new codebook that defined 12 categories (see Table 4.12).

In addition to the randomly selected 381-comment sample, I chose another 40 comments from the remaining 7,248 comments. These were saved in an Excel spreadsheet with comments in rows and categories in columns. Burns and I coded the comments independently on two separate Excel spreadsheets, and then met virtually to resolve conflicts. Inter-rater agreement was 92.5%. Adjustments were made to category definitions and a revised codebook was developed for coder training.

The 12 coding categories were entered as another set of descriptors in the Dedoose database, where I coded message thoughts for the entire sample set. The 381-comment sample set was downloaded from Dedoose and saved in an Excel spreadsheet in the manner described above. The spreadsheet was emailed to Crawford, who coded 40/381 comments (approximately 10%) so inter-rater agreement could be reported.

Recall that Creswell (2007) emphasized finding interconnected themes is necessary for rigorous research, making data analysis not a validation effort, but rather an exercise in leveraging grounded creation of new theoretical insights. When evaluating results from Chapter 4, I employed a grounded practical theory approach, which builds on elements of Glaser and Strauss' traditional grounded theory. The goal of grounded practical theory is to construct practical theory, rather than social-scientific theory. "The grounded practical theory method moves back and forth between interpretive empirical studies of particular communicative practices and an evolving normative model or 'rational reconstruction' that conceptualizes values and principles already partly implicit in those practices" (Craig & Tracy, 2009, p. 64). In the context of the retrospective explanatory case study research presented in this dissertation, grounded practical theory aided in categorizing factors that affected the use of the PubMed Commons forum. These in turn were used to develop the forum innovation agility model introduced in Chapter 5, which maps a strategic plan forum organizers might use when introducing, implementing, and managing a commenting forum similar to PubMed Commons.

3.6.3 Social Network Analysis

Numeric representations of commenters and their commenting information/ communication behavior were transformed from spreadsheet format into statistical computations using UCINET web-based software. Visual social relationship maps (i.e., sociograms) were created using the Gephi visualization tool in NetDraw, which is part of the

UCINET program. Both types of results (numeric and visual) are shown in Chapter 5 and are accompanied by a theoretical analysis that describes the significance of the findings.

Sociograms indicated individual member positioning in the network and described group characteristics based on proximity. Clusters and structural holes were highlighted, as were network roles (e.g., connector, broker, gatekeeper, isolate, influencer). The socio-centric (i.e., overall) network was examined for density, connectedness, and scattering of groups. On an individual level, commenters with high centrality degrees were placed in the middle of clusters as a consequence of normalizing the network. Positions indicated a commenter's ability to influence commenting on the forum. The SNA provided a snapshot of the network on the day it was discontinued, showing which commenters had been connected to others and characterizing how the network might have affected the ways in which PubMed Commons functioned as an online forum.

3.6.4 Application of Historical Data

Iterative analysis of historical data about PubMed Commons (e.g., announcement introducing the forum that appeared in *Nature,* NCBI Insights, and the NIH Director's Blog; posts on the PubMed Commons Blog; poster presentations; YouTube video recordings of PubMed Commons organizers talking about the forum) occurred concurrently with data analyses outlined in this chapter. Review of historical data continued to the point of saturation (i.e., no new categories or themes emerged), which was a signal that this dataset was complete.

All data were analyzed inductively, using content and thematic analysis. This contributed to the grounded practical theory analysis. The ongoing process provided an evolving understanding to answer my research questions and aided in the development the forum innovation agility model. Excerpts from historical data resources serve as evidence throughout my discussion and recommendations sections in Chapter 5. Cumulative analysis of the data overwhelmingly pointed to using Everett Rogers' Diffusion of Innovation Theory to explain the PubMed Commons phenomenon. The discussion in Chapter 5 is informed by my knowledge and project management consulting experiences, which I relied on to make recommendations for organizations similar to NCBI who are wondering about the costs and benefits of introducing, implementing, and managing a forum like PubMed Commons.

Netlytics. Web analytics using the Netlytics SaaS included a sociogram showing diffusion of tweets and retweets about PubMed Commons that spread throughout the Twitter network. Automated word clouds demonstrated frequently used words. Statistics were reported for the name network (i.e., who mentions whom) and chain network (i.e., who replies to whom) during an early- and late-stage of the PubMed Commons forum lifecycle.

3.7 Chapter 3 Summary

Explanatory case study research requires a multifaceted investigation using mixed methods. This chapter detailed the methodology I used to answer my overarching question: How did PubMed Commons function as an online forum for posting comments about published articles? I provided thorough descriptions of data collection and data analyses. A CMDA was executed in two parts: 1) quantitative content analysis to discover how often and in what ways did stakeholders participate in PubMed Commons and 2) qualitative content analysis to discover what was the subject matter of comments posted on PubMed Commons. A social network analysis was performed to examine what types of communities formed on PubMed Commons; and a systematized search for historical information about PubMed Commons exposed factors that contributed to or limited scientific discourse on PubMed Commons. A Dedoose database was created to discover characteristics of forum commenters. I also explained grounded theory procedures for developing categories and discovering interrelationships among multifarious factors.

This chapter explained how data analyses and inductive reasoning offered a holistic investigation of the complexities faced by PubMed Commons organizers who introduced PubMed Commons as a forum for hosting open discourse about published articles.

CHAPTER 4

RESULTS

The explanatory case study research detailed in this dissertation explored the U.S. National Library of Medicine's (NLM) PubMed Commons forum—a commenting system embedded within the PubMed search engine platform. This chapter presents data collection and analysis in three broad sections. The first section details the process for data collection, the way in which datasets were organized, and various tools used for data analysis. The second section addresses the overarching research question (RQ1): How did PubMed Commons function as an online forum for posting comments about published articles? Results are reported in four subsections: RQ 1a—extent to which adopters participated as members of PubMed Commons; RQ 1b—characteristics of forum commenters; RQ 1c—subject matter of comments posted on PubMed Commons; and RQ 1d—types of communities formed on PubMed Commons. A third section of Chapter 4 answers RQ2 and examines factors that contributed to or limited the scientific discourse on PubMed Commons. Taken as a whole, the multiple sources of data and their analysis provide a comprehensive overview of PubMed Commons.

4.1 PubMed Commons Data Collection

Data collection for this dissertation first began in 2014 for a pilot study my colleague Shelly Burns and I completed, and then presented at the 2015 National Conference on Health Communication, Marketing, and Media sponsored by the Centers for Disease Control and Prevention (see Appendix D). PubMed Commons remained a topic of interest, a focus of investigation, a source for data collection, and a reason for invited presentations from 2014 to

2019. The importance of investigating PubMed Commons became obvious on February 1, 2018, when an NCBI blog post announced that PubMed Commons was to be discontinued: "comments on articles indexed in PubMed will continue to be visible on PubMed and PubMed Commons through March 3, 2018, after which time they will be available for download from NCBI's website" (NCBI, 2018, para. 2). Hence, the collection of a final dataset for this dissertation became more immediate.

4.1.1 Data Sources

4.1.1.1 PubMed Search for PubMed Commons Records

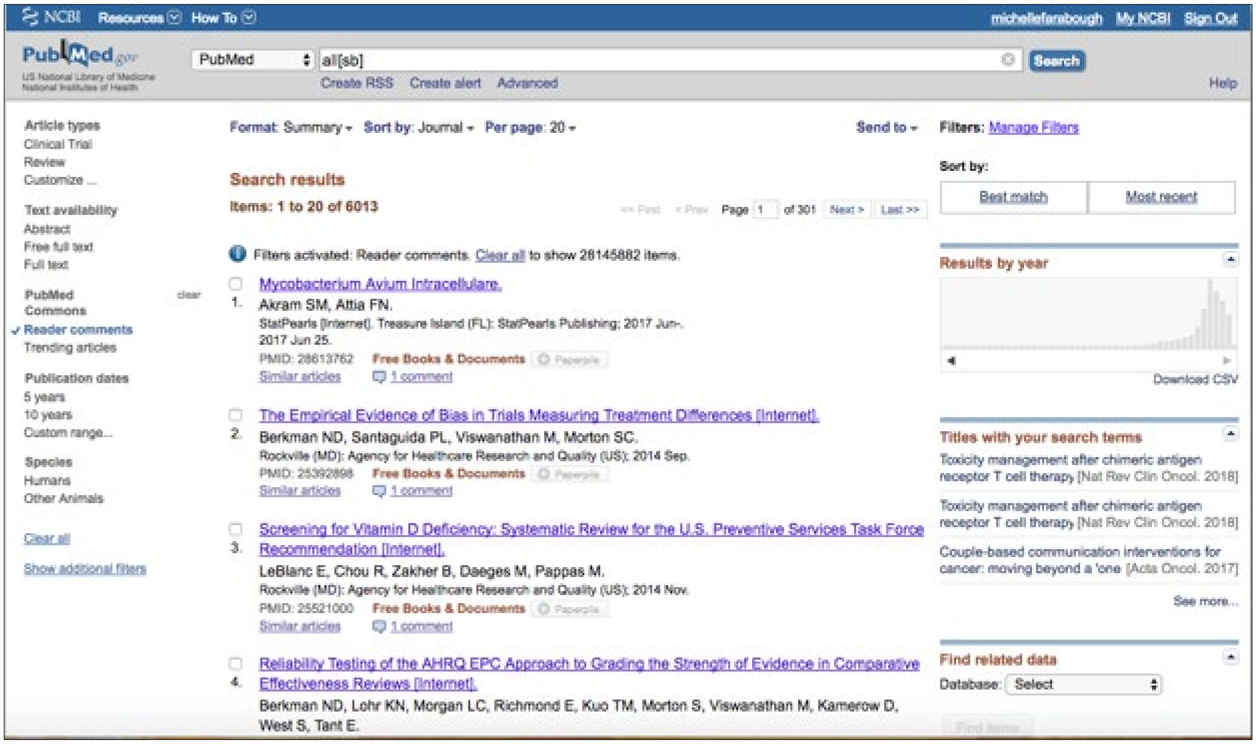
I conducted a PubMed search on February 18, 2018, by entering "all[sb]" into the PubMed search box and activated the "Reader comments" filter (see left column of Figure 4.1). The query returned 6,013 records, which is the number of PubMed-indexed articles that received at least one reader comment (see "Search results" in Figure 4.1).

The resulting PubMed search records were then sorted and downloaded into three comma-separated values (CSV) files: 1) "Most Recent," 2) "Journal," and 3) "First Author" (see Figure 4.2) This process aided the quantitative content analysis described in Section 4.5, which provides forum-, journal-, and commenter-level analyses.

The CSV files from the three data sets were saved as Excel spreadsheets, which in turn were copied and sorted in a variety of ways to address various identified categories of interest for each RQ. Columns were added to some of the spreadsheets to aid in a more granular quantitative content analysis. Appendix E shows examples of CSV and Excel files.

Figure 4.1

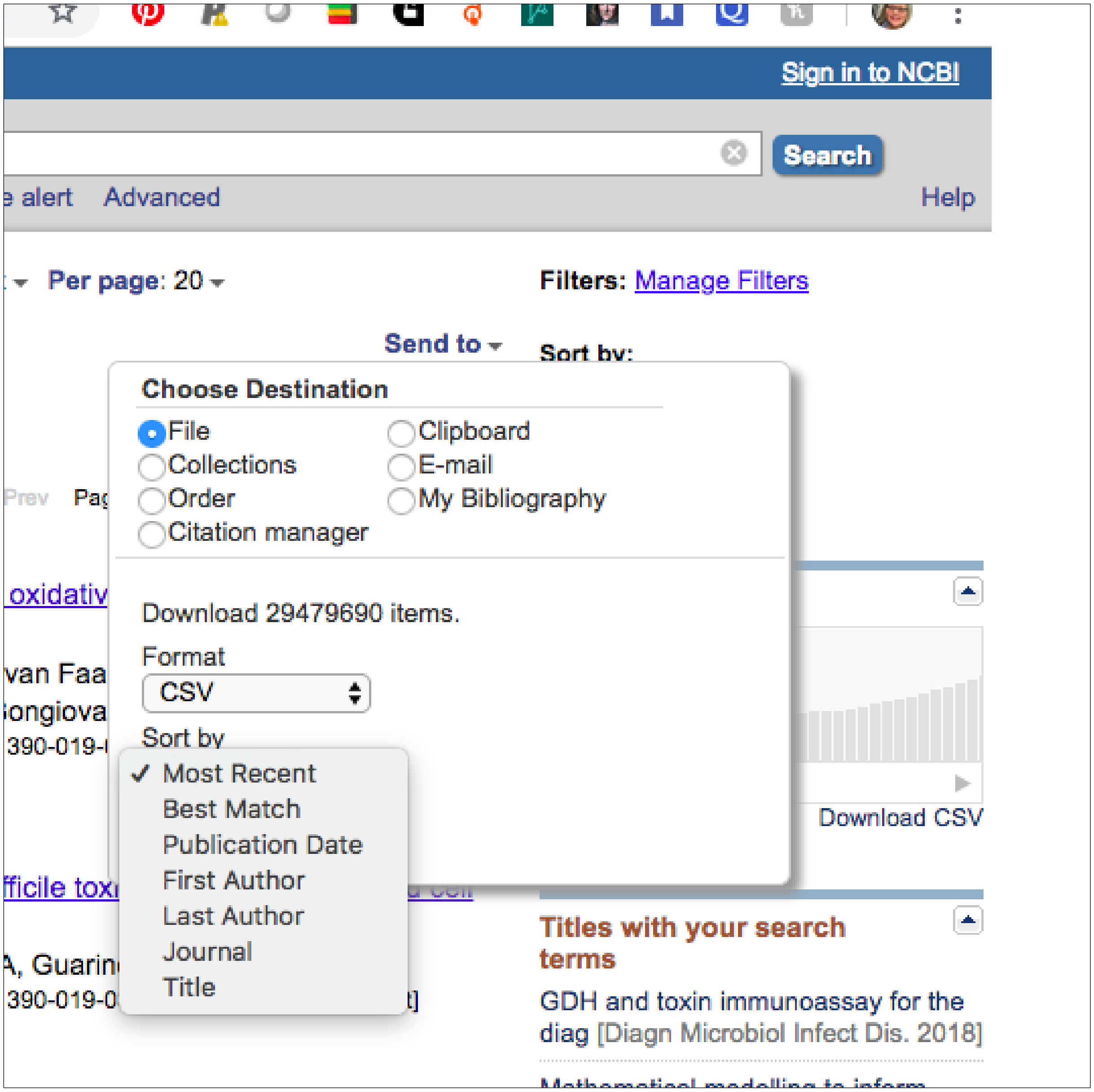
PubMed Search Results for All Indexed Articles with "Reader Comments"



To preserve the formatting integrity of records that appeared on the PubMed GUI following my PubMed search for "all[sb]," I copied and then pasted all search results onto an MS Word document (see Figure 4.3). The document was formatted in 10 pt. Calibri and filled 1,043 pages. Total word count was 354,241. This process captured and archived commenter name; date and time of post; if the post had been edited or deleted by the user or removed by the moderator; the number of people who found the comment helpful/not helpful; and the flow of comments (i.e., threaded non-related comments or nested replies). Copies of this master document were made, reorganized according to number of comments posted (e.g., 2, 3, 4, 5+) on any given PubMed-indexed article, and then saved for various data analyses.

Figure 4.2

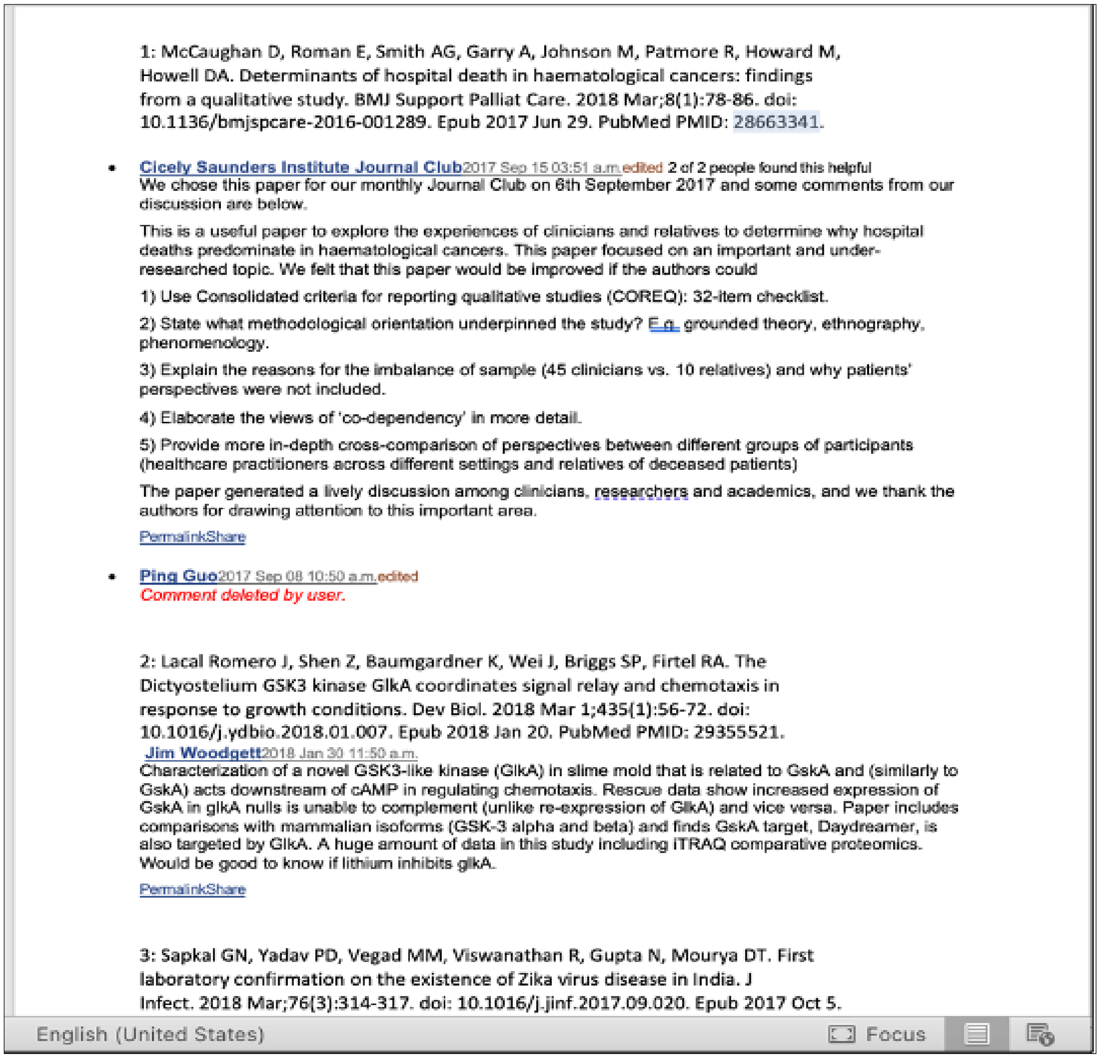
Sort Options for CSV File Indexed Articles with "Reader Comments"



Screenshot showing sort options for downloading query results for all indexed articles with "Reader comments."

Figure 4.3

PubMed Commons Dataset Copied and Pasted from PubMed GUI



1: McCaughan D, Roman E, Smith AG, Garry A, Johnson M, Patmore R, Howard M, Howell DA. Determinants of hospital death in haematological cancers: findings from a qualitative study. BMJ Support Palliat Care. 2018 Mar;8(1):78-86. doi: 10.1136/bmjspcare-2016-001289. Epub 2017 Jun 29. PubMed PMID: 28663341.

- Cicely Saunders Institute Journal Club2017 Sep 15 03:51 a.m.edited 2 of 2 people found this helpful
  We chose this paper for our monthly Journal Club on 6th September 2017 and some comments from our discussion are below.

  This is a useful paper to explore the experiences of clinicians and relatives to determine why hospital deaths predominate in haematological cancers. This paper focused on an important and under-researched topic. We felt that this paper would be improved if the authors could

  1) Use Consolidated criteria for reporting qualitative studies (COREQ): 32-item checklist.

  2) State what methodological orientation underpinned the study? E.g. grounded theory, ethnography, phenomenology.

  3) Explain the reasons for the imbalance of sample (45 clinicians vs. 10 relatives) and why patients' perspectives were not included.

  4) Elaborate the views of 'co-dependency' in more detail.

  5) Provide more in-depth cross-comparison of perspectives between different groups of participants (healthcare practitioners across different settings and relatives of deceased patients)

  The paper generated a lively discussion among clinicians, researchers and academics, and we thank the authors for drawing attention to this important area.

  PermalinkShare

- Ping Guo2017 Sep 08 10:50 a.m.edited
  Comment deleted by user.

2: Lacal Romero J, Shen Z, Baumgardner K, Wei J, Briggs SP, Firtel RA. The Dictyostelium GSK3 kinase GlkA coordinates signal relay and chemotaxis in response to growth conditions. Dev Biol. 2018 Mar 1;435(1):56-72. doi: 10.1016/j.ydbio.2018.01.007. Epub 2018 Jan 20. PubMed PMID: 29355521.

Jim Woodgett2018 Jan 30 11:50 a.m.
Characterization of a novel GSK3-like kinase (GlkA) in slime mold that is related to GskA and (similarly to GskA) acts downstream of cAMP in regulating chemotaxis. Rescue data show increased expression of GskA in glkA nulls is unable to complement (unlike re-expression of GlkA) and vice versa. Paper includes comparisons with mammalian isoforms (GSK-3 alpha and beta) and finds GskA target, Daydreamer, is also targeted by GlkA. A huge amount of data in this study including iTRAQ comparative proteomics. Would be good to know if lithium inhibits glkA.

PermalinkShare

3: Sapkal GN, Yadav PD, Vegad MM, Viswanathan R, Gupta N, Mourya DT. First laboratory confirmation on the existence of Zika virus disease in India. J Infect. 2018 Mar;76(3):314-317. doi: 10.1016/j.jinf.2017.09.020. Epub 2017 Oct 5.

English (United States) Focus

Screenshot of master MS Word document generated from copying/pasting results on the PubMed GUI following a search for "all[sb]" with the "Reader comments" filter activated.

4.1.1.2 Accessing the NCBI FTP Site "commons_archive.csv" File

The full dataset of 7,629 PubMed Commons comments was downloaded from the NCBI "FTP site" on March 3, 2018—the last day the forum was visible on the PubMed GUI. Total number of comments reflects the fact that some articles received two or more comments. As of Sept. 22, 2021, interested readers can access the "commons_archive.csv" file from the "FTP site” hyperlink at https://www.ncbi.nlm.nih.gov/pubmedcommons/ (see Figure 4.4). Current NCBI webpages with active hyperlinks to PubMed Commons-related words, phrases, or images are redirected to this webpage. Figure 4.5 shows hyperlinks on the FTP site.

Figure 4.4

Current NCBI Webpage Hyperlink from PubMed Commons References

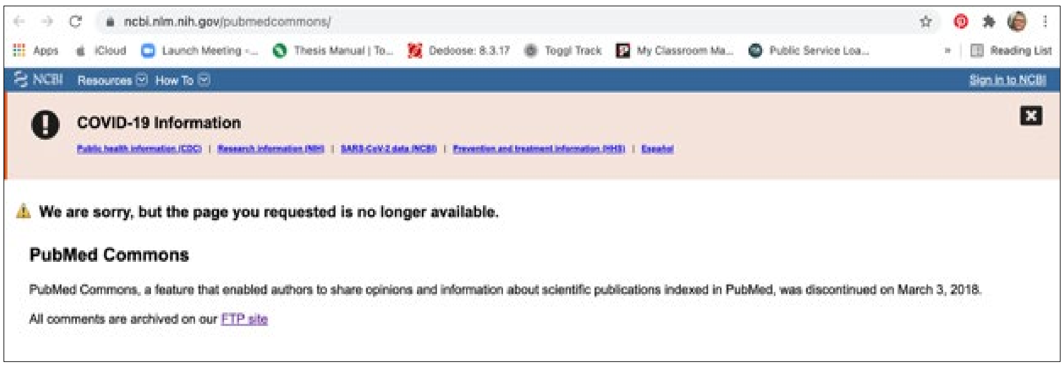


Figure 4.5

NCBI FTP Site with Link to PubMed Commons Archive

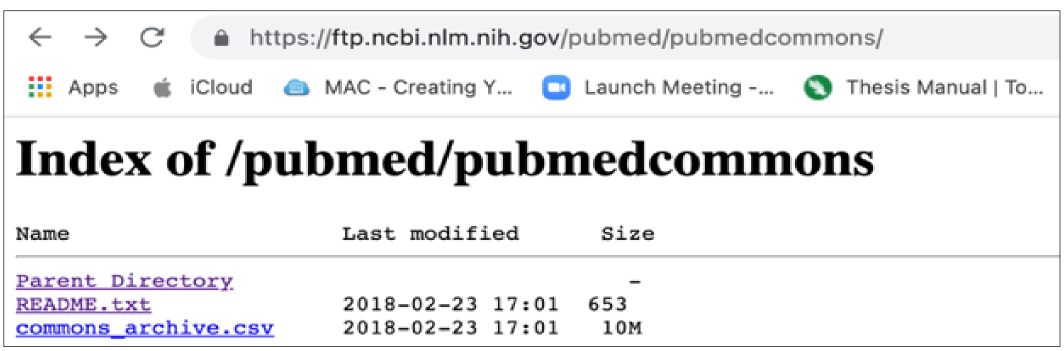


Figure 4.6 shows that captured data in the commons_archive.csv file includes only “Commentid”, “Pubmedid”, [commenter] “FirstName”, [commenter] “LastName”, and [post] “Content.” Unfortunately, necessary data for fully studying the PubMed Commons forum (e.g., threaded/nested comments, helpful ratings) is no longer available (see Figure 4.7).

Figure 4.6

Download of "commons_archive.csv" File

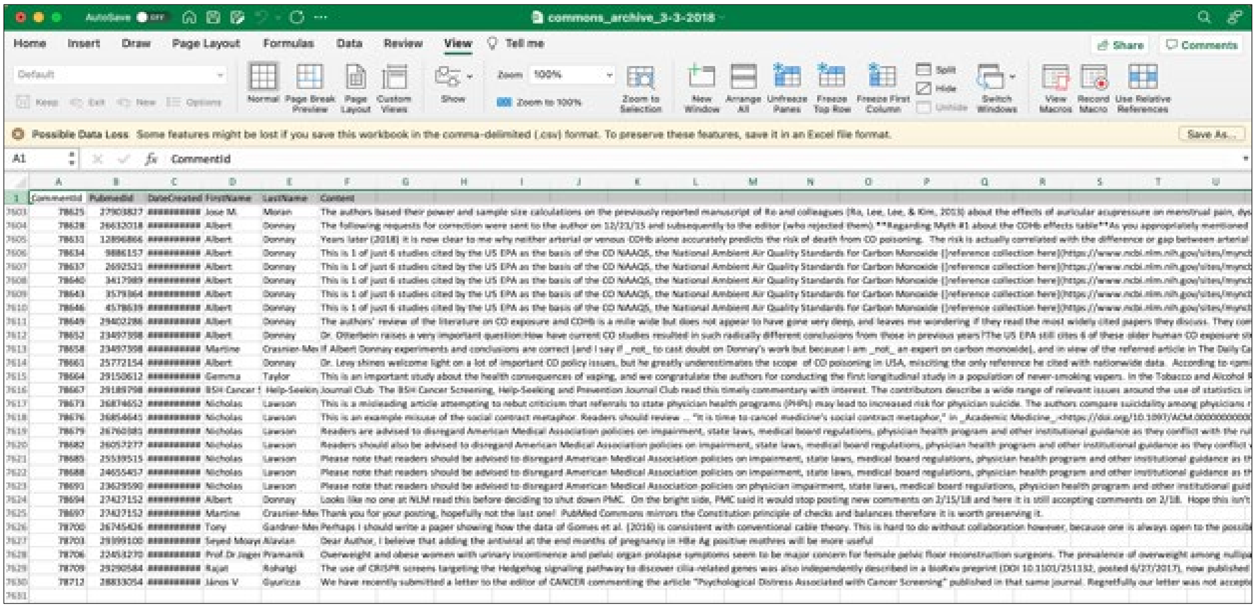


Figure 4.7

Data Showing Threaded Comments and Helpful Ratings

33: Pusapati GV, Kong JH, Patel BB, Krishnan A, Sagner A, Kinnebrew M, Briscoe J, Aravind L, Rohatgi R. CRISPR Screens Uncover Genes that Regulate Target Cell Sensitivity to the Morphogen Sonic Hedgehog. Dev Cell. 2018 Jan 8;44(1):113-129.e8. doi: 10.1016/j.devcel.2017.12.003. Epub 2017 Dec 28. Erratum in: Dev Cell. 2018 Jan 22;44(2):271. PubMed PMID: 29290584; PubMed Central PMCID: PMC5792066.

- Rajat Rohatgi2018 Feb 20 00:27 a.m. (yesterday)
  The use of CRISPR screens targeting the Hedgehog signaling pathway to discover cilia-related genes was also independently described in a bioRxiv preprint (DOI 10.1101/251132, posted 6/27/2017), now published in Nature Genetics (DOI 10.1038/s41588-018-0054-7)
  PermalinkShare

- Pete Monk2018 Jan 02 07:03 a.m. 3 of 3 people found this helpful
  Atthog is not a tetraspanin protein. Tetraspanins (TSPAN) are a specific gene family (https://www.genenames.org/cgi-bin/genefamilies/set/768). Atthog is unrelated to the TSPAN family but has 4 putative transmembrane domains; such proteins are known generically as tetraspans.
  PermalinkShare

  - Rajat Rohatgi2018 Jan 03 11:11 a.m. 3 of 3 people found this helpful
    Hi Pete, You are absolutely correct-- Atthog is a tetraspan protein, but does not belong to the tetraspanin family. Atthog is related to the claudin-like group of tetraspan proteins, but the term "tetraspanin" is specific to a different family. Thank you for bringing this error in nomenclature to our attention so promptly. We apologize for any confusion this may have caused. A correction is forthcoming.

My notions about losing such important details from the PubMed GUI were confirmed by others who had been investigating the forum. Following an initial data analysis of the commons_archive.csv file, bioinformatician Neil Saunders (2018) posted:

> There are likely to be other elements of interest in comment lists that we have not analysed. For example, some comments are nested in reply to previous comments but this is not captured by the CSS selectors in the current Ruby code. It would be of interest to see how many users reply to a comment, versus 'replying' using a top-level comment. It might also be interesting to analyse comment text using e.g., sentiment analysis. (paras. 19 and 20)

4.1.1.3 Systematized Search for Historical Information about PubMed Commons

From the onset of studying PubMed Commons in 2014, it was obvious that an ongoing environmental scan of the forum and its documented public record were important. I performed a systematized search every six months using the search term "PubMed Commons," looking for published literature and presentations focused on the forum, as well as information in grey literature and/or social media, including personal blogs that mentioned the forum. Searches were conducted in the Academic Search Complete, CINAHL, ERIC, JSTOR, ABI/INFORM, and Communication and Mass Media Complete databases, as well as on the PubMed, Google Scholar, and Google search engines. Each month I also visited NCBI-sponsored media outlets (e.g., PubMed Commons Blog*,* NCBI Insights*,* NLM digital newsletters/webpages, PubMed Commons' Twitter page). URLs, text, screenshots, PDFs, articles, and other files from the ongoing environmental scan of PubMed Commons were saved on my laptop in a "PubMed Commons General Info" folder.

4.1.1.4 Key Informant Interview and Personal Communication

In late Oct. 2016, I attended an NLM-hosted medical librarian training seminar in

Bethesda, MD. Ahead of the training, I arranged an interview with PubMed Commons editors Hilda Bastian and Melissa Vaught at their office in the NLM building on the National Institutes of Health (NIH) campus. The information they shared was invaluable for learning more about the forum's history and management. My colleague Julia Crawford (JC) took detailed notes, which enabled me to actively engage in the conversation without distraction. Interview questions are available in Appendix F.

4.2 Organizing PubMed Commons Data for Analyses

Table 4.1

Dara Sources for Analysis

Data Description	Data Format	Data Source	Data Use
PubMed "all[sb]" search	csv imported into MS Excel	PubMed search February 18, 2018	Article-, journal-, commenter-level overview
PubMed "all[sb]" search results on PubMed GUI	MS Word text file copied/pasted from PubMed GUI	PubMed search February 18, 2018	Edited, deleted, moderated, removed comments
commons_archives. csv file	csv imported into MS Excel	NCBI FTP site March 3, 2018	PMIDs
commons_archives. csv file	DOS Matrix	NCBI FTP site March 3, 2018	UCINET SNA
Commenter demographics (e.g., country of employment, university/organization affiliation, gender-if unknown,	Website text copy/paste into Dedoose database field responses	Google March 11, 2019 thru July 14, 2019	Commenter descriptors for Dedoose database
Article Altmetric scores	Website text copy/paste into Dedoose database field responses	Altmetric's Dimensions Discover database	

(table continues)

Data Description	Data Format	Data Source	Data Use
Systematized search; environmental scan for "PubMed Commons"	URLs for search results and text/graphics on hyperlinked webpages, Article and poster PDFs,	Google/Google Scholar search results, Academic databases, Twitter, PubMed Commons and NCBI Insights blogs, PubMed GUI	Background information and evidence for dissertation discussion, conclusions, and future research
PubMed Commons Twitter Account posts	Sociogram; Graphs; Text files	Netlytics data	SNA of commenters; Sentiment analysis of Twitter comments; PubMed Commons users' perceptions

Explanatory case study research of the online PubMed Commons forum is intended to serve as an agile and reflexive method for capturing emergent data about the large-scale post-publication social commenting phenomenon that was initiated by the NIH and hosted on the NLM's PubMed search engine. Since there has been little empirical research about this or a similar platform, this dissertation details how data was organized and analyzed to provide a template for similar research. To answer questions about how PubMed Commons functioned as a place for scientific discourse, collected data was organized and analyzed under different foci for specific purposes (e.g., perform a social network analysis, build a Dedoose database).

Table 4.1 details collected data, format, source, and use for results that are reported in this chapter. Organizing the data in such a fashion facilitated my ability to provide a comprehensive overview of forum use and the associated information behavior.

4.3 Tools for Analyses

During data collection, it became clear that studying PubMed Commons was complex. Explanatory case study research offered a guide for discovering causal links to explain the forum’s activity and its discontinuation. Accordingly, I accepted there would be no pre-determined outcome of my research and that data analyses should look at the phenomenon from many angles (Creswell, 2014; Yin, 2014). This dissertation uses a mixed methods approach, leveraging the best of quantitative and qualitative inquiry. I used a variety of methods and tools to answer my research questions, report my findings, and build a theory about “what happened on the way to the forum?”

4.3.1 Manual Processing

A purely quantitative approach to data collection and analysis is often criticized for failing to recognize important aspects of human lives (McCracken, 1988). While having researcher and phenomenon independent of one another aims to achieve objectivity, I maintain that my personal intimacy with the PubMed Commons data (e.g., manual counting, first-hand reading of all comments) and the iterative evaluation of all datasets resulted in contextualized knowledge about the PubMed Commons (Spencer, Pryce, & Walsh, 2014). For example, rather than utilize an open-source programming tool like Ruby code or developing other strings of code to perform automated data analyses (Saunders, 2018), I organized, sorted, counted, and manually calculated and analyzed CSV, Excel, and text files, looking for patterns and associations. Comparing my findings with Saunders’ automated data analyses were useful, however, for validating my own calculations.

4.3.2 UCINet

Social interaction data about PubMed Commons commenters (e.g., who commented on whose posts) was gathered by way of the text formatting of threaded/nested comments on the master MS Word document I generated from the PubMed GUI following a search for “all[sb]” with the “Reader comments” filter activated. Text formatting for threaded and nested comments was necessary for performing a social network analysis of the holistic forum and creating a visual map (i.e., sociogram) of the social networks that naturally occurred among commenting participants.

4.3.3 Dedoose

Twenty-five demographic and characteristic factors about commenters were entered into a Dedoose database that was created to analyze descriptive information about forum participants and make associations with their comments. Visual and statistical analyses of this data enhanced the understanding of PubMed Commons and those who made comments on the forum. Results validated quantitative content analysis results from manually processing the data.

4.4 RQ1: How PubMed Commons Functioned as an Online Forum

This section of Chapter 4 answers the overarching research question for this dissertation (i.e., How did PubMed Commons function as an online forum for posting comments about published articles?). Findings are reported in four subsections. The first provides a quantitative content analysis to answer RQ1a—to what extent did adopters participate as members of PubMed Commons? The second answers RQ1b, detailing characteristics of commenters on PubMed Commons. Attributes from documented participation are reported from a Dedoose database I built by inputting a data into a collection of 25 factors for a statistically significant sample of PubMed Commons commenters and their comments (381/7,629 comments or 5% of comments). The third subsection details results from a qualitative content analysis that answers RQ1c and identifies the subject matter of comments (i.e., comment content) posted on PubMed Commons. Each comment is categorized according to a validated set of codes developed from previous investigations of the forum. Furthermore, comment content is associated with descriptors from the Dedoose database to describe personas of commenters who authored comments posted on PubMed Commons. The fourth subsection reports a social network analysis that answers RQ1d by highlighting the types of communities that formed on PubMed Commons. Statistical analysis and sociogram visualizations identify key individuals in the PubMed Commons social network and detail their roles in connecting commenters and influencing the flow of commenting.

4.5 RQ1a: Extent to Which Stakeholders Participated in PubMed Commons

Once data had been downloaded and organized according to the process described above, a quantitative content analysis offered both a holistic and atomistic view about forum use, users, articles, journals, and comments. Results start to tell a story about participants who used the forum, patterns that emerged, forum management and oversight, technical functionality, and other factors that could have affected the forum's use.

4.5.1 Forum Overview

Following an initial soft launch in June 2013, PubMed Commons posting activity peaked in 2014, dropped nearly 30% in 2015, and rebounded slightly in 2016 before tapering off significantly until its discontinuation in 2018 (see Figure 4.8).

Figure 4.8

Number of PubMed Commons Comments by Year

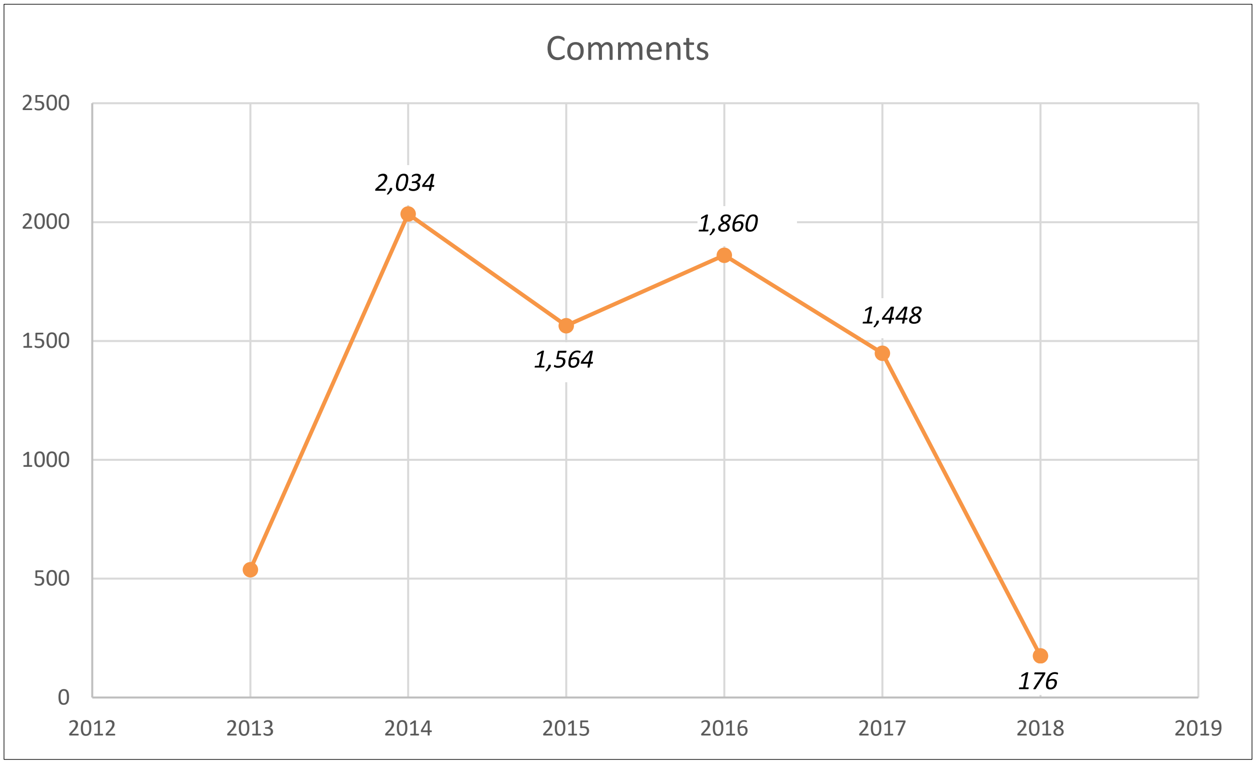


Table 4.2

PubMed Commons Quantitative Forum Overview

PubMed Commons Characteristics on March 3, 2018	Count
PubMed Commons comments	7,629
PubMed-indexed articles with comments	6,013
PubMed Commons comment word count	354,241
Unique journals with articles receiving comments	1,854
Unique individuals posting on PubMed Commons	1,551
Journal Club commenters	18
Deleted comments	236 (143 by user; 93 by moderator)
Number of PubMed-indexed articles receiving only one comment	5,210
Percent of PubMed-indexed articles receiving comments	.0002% of 28,145,882

Table 4.3

PubMed Commons Averaged Data

PubMed Commons Characteristic on March 3, 2018	Average
PubMed Commons comments per PubMed-indexed article	3.84
PubMed Commons comments per commenter	4.43
PubMed-indexed article per commenter	3.67
Deleted or Moderator Removed PubMed Commons comments per commenter	3.15
Word count per comment	82.17

A total of 7,629 comments were posted on the 6,013 PubMed-indexed articles that received comments. Less than two one-hundredths of a percent (.00027105) of the over 28 million PubMed-indexed articles (at the time of the forum's discontinuation) received a comment on PubMed Commons. Approximately 68% of the articles had only a single comment, and 3% of those were either deleted by the commenter or removed by the moderator. See Table 4.2 and 4.3 for a detailed numeric overview of the forum.

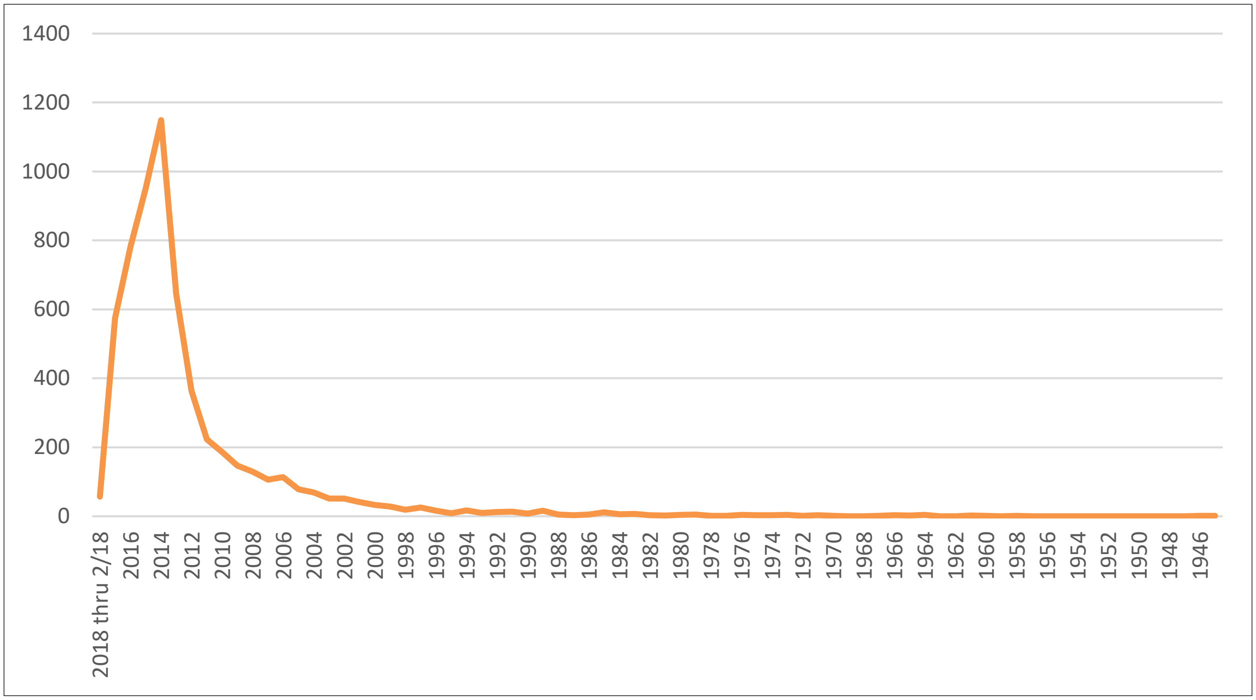
On Apr. 6, 2017, the PubMed Commons blog reported that thru March 31, 2017, the forum had 10,962 members of which 1,637 had commented. According to NCBI, 6,372 comments had been made on 5,078 PubMed-indexed publications (see Appendix A).

4.5.2 Article-Level Analysis

PubMed Commons comments were made on PubMed-indexed articles published as early as Aug. 25, 1945, and as late as Feb. 7, 2018 (see Figure 4.9). The graph below shows that most comments were made on more recently published articles. Articles published in 2014 received the most comments.

Figure 4.9

Distribution of PubMed Commons Comments by Publication Year



"Accidental Infant Suffocation" authored by W. H. Davison, published in the *British Medical Journal*, and indexed Aug. 25, 1945, was the earliest article receiving a comment. The single comment posted on the article abstract record gave a historical analysis of risk factors for infant death and refuted the author's original findings, informing that cause of death was not a prone sleeping position, but rather respiratory or Otis media. "T Cell Expression of C5a Receptor 2 Augments Murine Regulatory T Cell (T $_{\text{REG}}$) Generation and T $_{\text{REG}}$–Dependent Cardiac Allograft Survival" authored by D. A. Verghese, M. Demire, N. Chun, M. Friburg, P. Cravedi, I. Llaudo, T. M. Woodruff, P. Yadav, S. A. Lira, M. E. Medof, & P. S. Heeger" and indexed Feb. 7, 2018 ahead of the Mar. 15, 2018, publishing date in the *Journal of Immunology* was the final article receiving a comment. The single comment posted on the article record redirected readers to a

2009 article, which the commenter claimed was the first publication evidencing the role of C5aR2 as a scavenger or decoy receptor for C5a.

PubMed-indexed articles receiving eight or more comments are shown in Table 4.4. “When is Science ‘Ultimately Unreliable’?” authored by Michael Blatt and published Mar. 2016 in *Plant Physiology* received the most comments (33), although only six commenters were involved in the discussion. Blatt is the Regius professor of botany at the University of Glasgow’s Institute of Molecular, Cell and Systems Biology and editor of *Plant Physiology.* Commenters, including Blatt, discussed anonymity in post-publication peer review. Blatt chiefly debated Boris Barbour, a neuroscientist from France (École normale supérieure [ENS] Paris, 2021) and co-organizer of “PubPeer: The online Journal club,” where commenters can remain anonymous.

Another article authored by Blatt—Vigilante Science” published Oct. 2015 in *Plant Physiology*—received the second highest number of comments (26). Eight commenters, including Blatt and Barbour, kicked off their discussion about post-publication commenting anonymity. Table 4.4 shows that articles with a high number of comments rarely had more than a few commenters (see “Commenters” column in Table 4.4).

The PubMed-indexed article with the most helpful/not helpful comments (123) was “The Heroes of CRISPR” authored by Eric S. Lander and published Jan. 14, 2016, in *Cell.* Sixty individuals found Jennifer Doudna’s comment helpful (see Figure 4.10). Doudna is a Nobel Laureate in chemistry and Li Ka Shing Chancellor’s professor of biomedical science for University of California Berkeley Research (University of California Berkeley, n.d.). She is principal investigator at the Doudna Lab.

Table 4.4

Articles Receiving Eight or More Comments

PMID	Comments	Commenters	Title	First Author	Journal
26933091	33	6	When is science 'ultimately unreliable'?	Blatt, M. R.	*Plant Physiol*
26417050	26	8	Vigilante Science	Blatt, M. R.	*Plant Physiol*
25219520	22	4	Uncovering the hidden risk architecture of the schizophrenias: Confirmation in three independent genome-wide association studies	Arnedo, J.	*Am J Psychiatry*
24021304	16	11	Assessment of causality of individual adverse events following immunization (AEFI): A WHO took for global use	Tozzi, A. E.	*Vaccine*
25268438	13	5	Acupuncture for chronic Knee pain: A randomized clinical trial	Hinman, R. S.	*JAMA*
25554788	12	12	Cancer etiology. Variation in cancer risk among tissues can be explained by the number of stem cell divisions	Tomasetti, C.	*Science*
27620683	12	7	The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses	Ioannidis, J. P.	*Milbank Q.*
28396415	12	2	Variations in crowding, saccadic precision, and spatial localization reveal the shared topology of spatial vision	Greenwood, J. A.	*Proc Natl Acad Sci USA*
28029926	11	4	Fish oil-derived fatty acids in pregnancy and wheeze and asthma in offspring	Bisgaard, H.	*N Engl J Med*
28971835	11	2	The pump, the exchanger, and the holy spirit: Origins and 40-year evolution of ideas about the ouabain-Na pump endocrine system	Blaustein, M.	*Am J Physiol Cell Physiol*
27518691	9	5	Evidence, policy, and e-cigarettes	McKee, M.	*N Engl J Med*
12053565	9	1	New definitions of the concepts and terms ecosystem and biogeocenosis	*Ostroumov, S.	*Dokl Biol Sci*
26933091	33	6	When is science 'ultimately unreliable'?	Blatt, M. R.	*Plant Physiol*

(table continues)

PMID	Comments	Commenters	Title	First Author	Journal
25739399	9	4	Wikipedia and medicine: Quantifying readership, editors, and the significance of natural language	Heilman, J. M.	*J Med Internet Res*
23363640	8	4	Recovery from chronic fatigue syndrome after treatments given in the PACE trial	White, P.D.	*Psychol Med*
20143388	8	4	Check your cultures! A list of cross-contaminated or misidentified cell lines	*Capes-Davis, A.	*Int J Cancer*
21334061	8	5	Comparison of adaptive pacing therapy, cognitive behaviour therapy, graded exercise therapy, and specialist medical care for chronic fatigue syndrome (PACE): A randomized trial	White, P.D.	*Lancet*
26129895	8	4	Demystifying the search button: A comprehensive PubMed search strategy for performing an exhaustive literature review	McKeever, L.	*JPEN J Parenter Enteral Nutr*
27693003	8	3	Electronic cigarettes increase endothelial progenitor cells in the blood of healthy volunteers	Antoniewicz, L.	*Atherosclerosis*
27934275	8	3	Flavoring compounds dominate toxic aldehyde production during e-cigarette vaping	Khlystov, A.	*Environ Sci Technol*
26524703	8	2	Moderate alcohol consumption is not associated with reduced all-cause mortality	Goulden, R.	*Am J Med*
28074888	8	2	A review of the carbohydrate-insulin model of obesity	Hall, K. D.	*Eur J Clin Nutr*
20877712	8	5	Seventy-five trials and eleven systematic reviews a day: How will we ever keep up?	*Bastian, H.	*PLoS Med*

* Top 20 PubMed Commons commenters with most comments

Figure 4.10

Most Helpful PubMed Commons Comment

Claudiu Bandea 2016 Jan 24 6:59 p.m. 1 of 10 people found this helpful
Comment removed by moderators.

Emmanuelle Charpentier 2016 Jan 19 5:09 p.m. 41 of 41 people found this helpful
I regret that the description of my and collaborators' contributions is incomplete and inaccurate. The author did not ask me to check statements regarding me or my lab. I did not see any part of this paper prior to its submission by the author. And the journal did not involve me in the review process.

PermalinkShare

JENNIFER DOUDNA 2016 Jan 17 10:31 p.m. 60 of 60 people found this helpful
From Cell editor: "…the author engaged in substantial fact checking directly with the relevant individuals."

However, the description of my lab's research and our interactions with other investigators is factually incorrect, was not checked by the author and was not agreed to by me prior to publication.

PermalinkShare

This article was mentioned in a comment by Donald Forsdyke 2016 Jul 21 2:55 p.m.
See: Viruses are a dominant driver of protein adaptation in mammals. [Elife. 2016.]

4.5.3 Journal-level Analysis

Knowing the journal titles and number of PubMed Commons comments posted on PubMed-indexed articles published in individual journals provides insights about which fields of research or topics of interest are more likely to stimulate scientific discourse and which journals have engaged readership. The distribution of comments over the number of journals publishing articles that received comments is shown in Figure 4.11. Less than half of the 1,854 unique journals published articles received two or more comments on PubMed Commons.

Table 4.5 provides details and metrics for journals that published PubMed-indexed articles receiving more than 20 reader comments. At the time of data collection, 12 journals (60%) on the Top 20 list were considered tier one (i.e., impact factor 12.001–999.99); seven (35%) were tier two (i.e., impact factor 6.001–12.0); six (30%) were tier three (i.e., impact factor 2.6–6.0); and one (5%) was tier four (i.e., impact factor 1–2.599). None were tier five (i.e., impact factor 0.001–0.999). The table indicates which journals are fully open access, which might influence scientific impact (Bjork & Solomon, 2012).

Figure 4.11

Breakdown of Number of Comments per Unique Journal

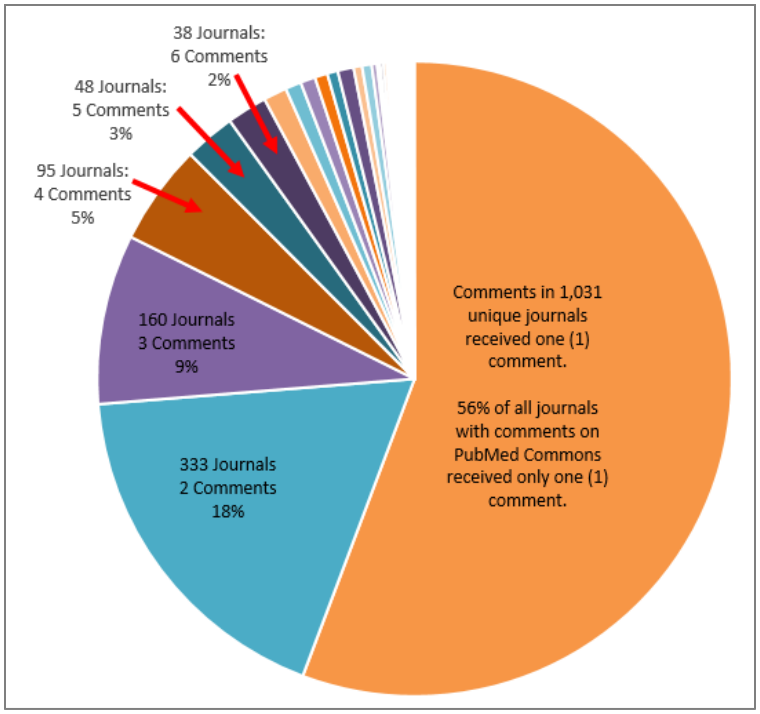


Table 4.5

Journals with Articles Receiving 20 or More Comments

Number of Comments	Journal Title	*Tier Designation[a]	Impact Factor (5-year)[a]	Discipline[a]	Rejection Rate (%)[b]	Open Access
207	*PLoS One†*	3	3.352	Multidisciplinary Sciences	50	Fully
156	*New England Journal of Medicine†*	1	67.513	Medicine, General & Internal	95	
121	*Nature†*	1	44.959	Multidisciplinary Sciences	92	
114	*Proceedings of the National Academy of Sciences of the United States of America†*	2	10.359	Multidisciplinary Sciences	83	

(table continues)

Number of Comments	Journal Title	*Tier Designation[a]	Impact Factor (5-year)[a]	Discipline[a]	Rejection Rate (%)[b]	Open Access
100	*JAMA†*	1	42.464	Medicine, General & Internal	88	
78	*Science†*	1	40.627	Multidisciplinary Sciences	84	
76	*Lancet†*	1	52.665	Medicine, General & Internal	95	
73	*Journal of Vision*	4	2.489	Ophthalmology	50	Fully
67	*Nucleic Acids Research†*	2	10.235	Biochemistry & Molecular Biology	67	Fully
58	*BMJ†*	1	23.562	Medicine, General & Internal	93	Fully
55	*JAMA Internal Medicine†*	1	17.84	Medicine, General & Internal	86	
54	*Cell†*	1	33.796	Biochemistry & Molecular Biology; Cell Biology	N/A	
53	*Scientific Reports*	3	4.609	Multidisciplinary Sciences	41	Fully
52	*Cochrane Database Systematic Reviews†*	2	7.669	Medicine, General & Internal	N/A	
46	*Journal of Biological Chemistry†*	3	4.254	Biochemistry & Molecular Biology	57	Fully
38	*Bioinformatics†*	2	8.561	Biochemical Research Methods; Biotechnology & Applied Microbiology	65	
38	*Nature Communications†*	1	13.691	Multidisciplinary Sciences	83	Fully
31	*Annals of Internal Medicine†*	1	18.726	Medicine, General & Internal	93	
28	*Pediatrics†*	2	6.442	Pediatrics	83	
27	*Journal of Neuroscience†*	2	6.518	Neurosciences	79	
25	*Cancer Research*	2	9.578	Oncology	80	
24	*Circulation*	1	17.902	Cardiac & Cardiovascular Systems; Peripheral Vascular Disease	92	

(table continues)

Number of Comments	Journal Title	*Tier Designation[a]	Impact Factor (5-year)[a]	Discipline[a]	Rejection Rate (%)[b]	Open Access
24	*Journal of Clinical Epidemiology*	3	5.185	Health Care Sciences & Services; Public; Environmental & Occupational Health	72	
23	*BMC Bioinformatics*	3	3.114	Biochemcial Research Methods; Biotechnology & Applied Bicrobiology; Mathematical & Computations Biology	45	Fully
23	*Nature Neuroscience*	1	19.188	Neurosciences	N/A	
21	*Annals of Emergency Medicine*	3	5.441	Emergency Medicine	85	

*Tiers: 1 (IF 12.001–999.99); 2 (IF 6.001–12.0); 3 (IF2.6–6.0); 4 (IF 1–2.599); 5 (IF 0.001–0.999). † Lane (2016) Top journals publishing articles with the greatest number of comments on PubMed Commons. a Web of Science Incites Journal Citation Reports. bPubsHub Journals & Congresses

Table 4.6 informs about journals that published PubMed-indexed articles receiving fewer than 20 comments. The high incidence of single comments is notable.

Table 4.6

Journals with Articles Receiving Fewer Than 20 Comments

Number of Comments	Number of Unique Journals Receiving Comments	Percent (%) of Journals with PubMed-indexed Articles that Received a Comment
19	3	< 1
18	3	< 1
17	2	< 1
16	1	< 1
15	5	< 1
14	9	< 1
13	8	< 1
12	15	< 1

(table continues)

Number of Comments	Number of Unique Journals Receiving Comments	Percent (%) of Journals with PubMed-indexed Articles that Received a Comment
11	10	< 1
10	12	< 1
9	14	< 1
8	15	< 1
7	22	1
6	38	2
5	48	3
4	95	5
3	160	9
2	333	18
1	1031	56

4.5.4 Comment-Level Analysis

Detailing the number of unique individual and journal club commenters demonstrates the reach PubMed Commons had for attracting scientific discourse. Reporting the number and percentage of comments posted by each unique commenter provides insights about commenters' level of engagement (see Figure 4.12). The PubMed Commons forum was characterized by 1,551 unique commenters: 1,533 individuals and 18 journal clubs. The most frequent number of comments was one.

The first PubMed Commons comment was made June 12, 2013, during a closed pilot testing phase. Open piloting began in October 2013. Further evidence of commenter participation on the forum can be measured by the number of comments posted on any given PubMed-indexed article (i.e., comments in thread) and changes in posting behavior over time (see Table 4.7). Numbers echo the same patterns seen in Figure 4.8 above, which shows an uptake of the forum in 2014 and a sharp decline in activity in 2017.

Figure 4.12

Breakdown of Number of PubMed Commons Comments per Commenter

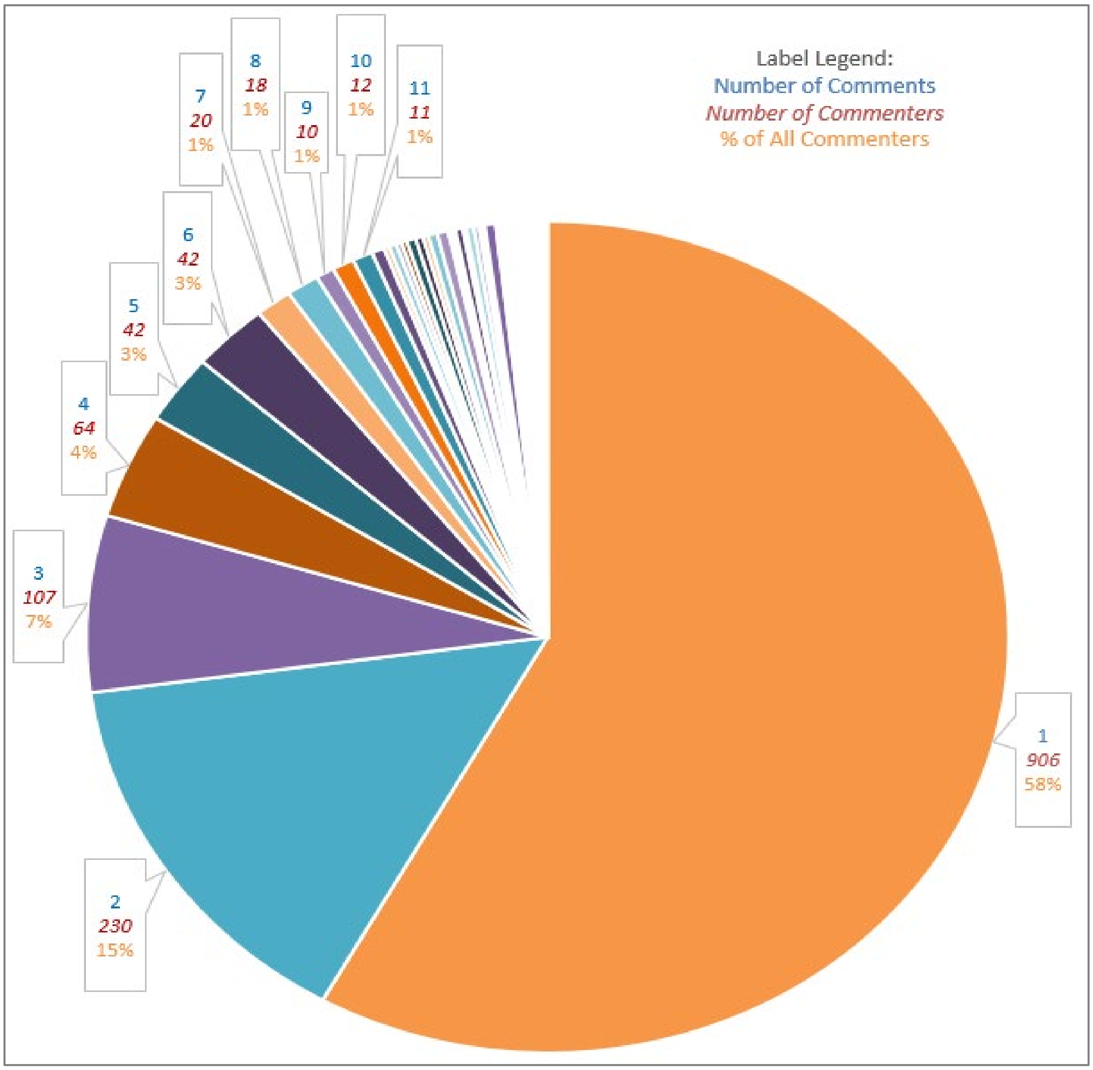


Table 4.7

Articles Receiving 3 or More Comments by Year

Number of Comments	Number of Articles	2013	2014	2015	2016	2017	2018
3	139	15	81	83	133	96	9
4	60	12	45	34	101	48	4
5	33	12	12	47	53	35	6

(table continues)

Number of Comments	Number of Articles	2013	2014	2015	2016	2017	2018
6	18	0	29	21	31	27	0
7	9	2	19	7	20	17	0
8	10	8	15	7	19	25	6
9	6	0	9	21	15	9	0
10	1	2	2	2	3	1	0
11	1	0	0	0	0	11	0
12	2	0	0	12	11	1	0
13	1	0	13	0	0	0	0
14	2	0	14	0	1	6	7
16	1	0	2	11	3	0	0
17	1	0	17	0	0	0	0
19	1	0	0	19	0	0	0
27	1	0	25	1	1	0	0
33	1	0	0	0	33	0	0

Table 4.8

Top 20 Commenters Who Commented Most Frequently

Commenter	Comments	Articles	Commenter interactions	Deleted by User	Removed by Moderators	Comment Context (See Section #)
Maniatis, Lydia*	255	141	17	4	3	Discussion
Keller, David*	234	160	10	0	0	Discussion
Kindlon, Tom*	165	105	11	3	0	Discussion
Goldacre, Ben	145	145	5	145	0	Watchdog Housekeeping
Oksvold, Morten	139	120	6	0	5	Watchdog
Hemila, Harri	137	134	3	1	2	Curation
Southan, Christopher*	117	114	15	8	0	Discussion
Wright, Kath	111	111	0	1	0	Self-promotion
Capes-Davis, Amanda*	89	85	7	2	0	PPPR

(table continues)

Commenter	Comments	Articles	Commenter interactions	Deleted by User	Removed by Moderators	Comment Context (See Section #)
Oransky, Ivan	88	86	8	1	0	Watchdog
Ostroumov, S A	81	54	5	0	0	Endorsement Curation
Cayley, Bill	78	77	3	1	0	Self-promotion
Cannell, John	73	71	1	2	9	Discussion Self-promotion
Bastian, Hilda*	71	60	51	0	0	Discussion
Ekins, Sean*	67	49	3	0	0	Curation Self-promotion
Radecki, Ryan	61	61	9	0	0	Discussion Self-promotion
Pechacek, Randi	59	59	2	0	0	Self-promotion
Bates, Clive	54	43	31	1	1	Critique
Forsdyke, Donald	53	49	9	0	1	PPPR
Eisen, Jonathan	50	46	1	0	0	Curation

* PubPeer Commenter. PPPR is an abbreviation for Post-publication Peer Review.

Learning more about the individuals who commented most frequently on PubMed Commons and characterizing their posting behavior provides insight about comment content and purpose. Close examination also informs about an individual's engagement in online scientific discourse, the community relationships they form, and shared interests with other commenters. See Table 4.8 for a list of commenters who posted most frequently. The narrative below the table explains the ways in which some individuals had similar commenting behavior.

Some commenters appeared more self-serving than others. Kath Wright's 111 comments on 111 articles were identical, promoting the use of the InterTASC Information Specialists' Sub-group Search Filters Resource. She was lead editor of the world-renown group until 2020 (InterTAC Information Specialists' Sub-Group [ISSG], 2021). None of Wright's posts were part of a threaded discussion. Bill Cayley, MD, posted 78 comments on 77 articles and promoted his evidence-based medicine “< = > Less Is More” blog, where he explains “simpler and better” and narrative medicine (BMJ Opinion, 2016). Cayley’s comments engaged three PubMed Commons commenters.

Other commenters appeared to be on a mission. Ben Goldacre’s posts were exclusively centered on correcting the published record of trial registry IDs associated with articles indexed on PubMed. His 145 comments on 145 articles garnered posts from five other commenters, each of whom thanked him for posting the correction (e.g., Joanna Hudson commented, “We thank Ben Goldacre and the opentrials.net project for bringing this to our attention. We have informed the journal of this anomaly and will ensure all future publications are cited with the correct trial registry ID”). Goldacre—psychiatrist, academic, writer, and broadcaster—is director of the University of Oxford DataLab (University of Oxford, 2021) and collaborator with Open Knowledge International on the OpenTrials (2021) project, which is an online database hosting information about the world’s clinical research trials.

One group of commenters frequently directed readers to other information resources. Harri Hemila’s 137 comments on 134 articles primarily redirected readers to resources for accessing publications, blog posts and comments on other media, and retraction notices. He interacted with three forum commenters. Hemila (n.d.), MD, PhD, teaches at the University of Helsinki, Finland and is a noted vitamin researcher. Sergei (“S.A.”) Ostroumov’s posting behavior (81 comments on 54 articles) was similar to Hemila’s in that he linked additional information found elsewhere. His comments were complimentary in nature (e.g., “At World Catalog, the paper was reviewed and rated as excellent”; “I think that the importance of microbial pollution will increase in future. Therefore, I consider this paper useful and relevant”).

Ostroumov, PhD, is a biological sciences researcher and highly prolific author who works at Moscow State University (n.d.). He is a member of the Russian Academy of Natural Sciences and section chair of the Moscow Society of Researchers of Nature, which was founded in 1805. Sean Ekins focused his 67 comments/49 articles on curating additional information about his work (e.g., updates and media attention) and recruiting collaborators. He also promoted startup companies he worked with. Ekins, PhD, Msc, Dsc, is founder and CEO of Collaborations Pharmaceutical, Inc. His career has centered on drug discovery and informatics. Ekins commented on articles with posts from three other commenters, including Christopher Southan—another top PubMed Commons commenter. Both were closely connected in the pharma industry.

John Cannell's comments were clearly focused on a single topic—autism spectrum disorder (ASD). He repeatedly shared his convictions that there is little communication among autism scientists. Each of his 73 comments on 71 articles posed rhetorical questions for authors to consider about autism etiology, particularly Vitamin D. Cannell is a psychiatrist and the founder of the Vitamin D Council (Blooming Wellness, 2018). One commenter focused his only PubMed Commons' post on validating Cannell's information, thanking him for his insightful comment, and agreeing that abnormalities in Vitamin D for the ASD population is significant and important.

Some commenters served as watchdogs of published literature. Ivan Oranksy's 88 comments on 86 articles exclusively warn of article retractions and corrections. His posts were part of forum discourse with eight other commenters. Oransky manages the *Retraction Watch* blog, which "track[s] retractions as a window into the scientific process" (Retraction Watch, n.d.). An illustrious writer and editor, he is currently Editor-in-Chief of *Spectrum* and a Distinguished Writer in Residence at New York University's Carter Journalism Institute where he teaches medical journalism. He is also the president of the Association of Health Care Journalism. Morten Oksvold's 139 comments on 120 articles also alerted readers about published "bad science," citing investigations, retractions, and formal scholarly communications. Oksvold, PhD researcher and author, works at the Oslo University Hospital (n.d.) where he studies and writes about laboratory medicine.

A handful of commenters wrote detailed comments that were carefully crafted literary contributions. This group included Lydia Maniatis, PhD psychology professor at University of North Carolina, (255 comments on 141 articles); David Keller, MD, FACP, former internal medicine doctor, (234 comments on 160 articles); Tom Kindlon, researcher and myalgia encephalomyelitis advocate, (165 comments on 105 articles); PubMed Commons editor Bastian (71 comments on 60 articles); and Southan, PhD member of the British Pharmacological Society, (117 comments on 114 articles). Their posts demonstrated respect for the forum as a vehicle for medical publishing. These commenters posted on several different articles and were participated in lively, online discourse.

Clive Bates' 54 comments on 43 articles were more critical in nature than the others penned by individuals in this elite group of frequent commenters. Bates—"The Counterfactual" blogger (2021), Director of Counterfactual Consulting Limited, and former Director General of the Dept. of Energy and Climate Change for the Welsh Government—is an analytical advocate. His PubMed Commons comments drew attention to his interpretations of articles' false claims, faulty conclusions, weaknesses, and conflicts of interest with research and author sponsorships.

Amanda Capes-Davis, MB, MS, BSc (Med), PhD, used her expertise as member of the International Cell Line Authentication Committee (ICLAC) and Founding Manager at CellBank Australia to advance post-publication peer review. She validated and critiqued 85 articles and posted 84 comments, focusing attention on cell culture practice, authentication testing, and improving approaches to misidentified cell lines (ResearchGate, 2021).

When examining information behavior of PubMed Commons commenters, I became interested in the relationships that were forming. Table 4.9 shows the list of commenters who interacted with 10 or more commenters. My iterative data analysis prompted a social network analysis of the forum.

Table 4.9

Commenters Who Interacted with 10 or More Commenters

Commenter	Other Commenters
Bastian, Hilda	47
Bates, Clive	23
Bishop, Dorothy	17
Maniatis, Lydia	17
Bramer, Wichor	15
Corcos, Daniel	15
Southan, Christopher	15
Lopez-Lazaro, Miguel	14
Detours, Vincent	13
Grant, William	13
Hegde, B M	12
Barbour, Boris	11
Brody, Jim	11
Burkitt, Mark	11

(table continues)

Commenter	Other Commenters
Ciulla, Michele	11
Coletta, Andrea	11
Delpierre, Cyril	11
Girish, Meenakshi	11
Kuznetsov, Vladimir	11
Oransky, Ivan	11
Girard, Marc	10
King, Paul	10
Malik, Akash	10
Puliyel, Paul	10
Teixeira da Silva, Jaimie	10
Tiwari, Lokesh	10

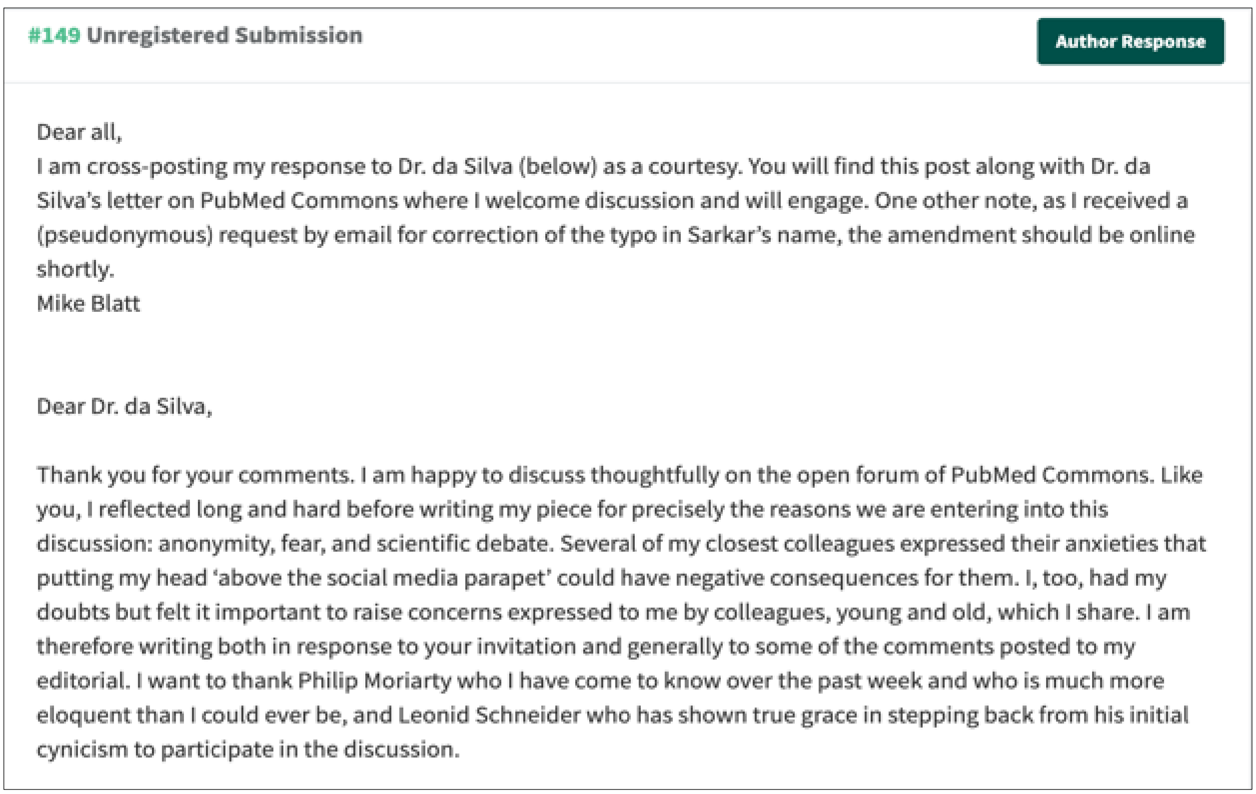
Top PubMed Commons commenter Maniatis was part of an engaged group of six commenters who posted on PubMed Commons most commented-on article—”When is Science ‘Ultimately Unreliable’?” PubPeer co-organizer Boris Barbour (9/33 comments), Mantiatis (5/33 comments), and article author Blatt (12/33 comments) held a heated, albeit respectful, discussion about anonymous post-publication peer review commenting. The discussion was also archived on PubPeer, which attracted an additional 18 anonymous comments on its website.

PubPeer is an online forum launched in 2012 that enables its users to anonymously discuss and review scientific research after publication. The online platform is hosted by the PubPeer Foundation (2021) with a goal “to improve the quality of scientific research by enabling innovative approaches for community interaction” (para. 1). Further information about PubPeer and anonymous commenting is presented in Chapter 5 as a factor affecting the use of PubMed Commons.

The same group who commented on the 'Ultimately Unreliable' article first began their discussion in 2015 on the PubMed-indexed article "Vigilante Science," which received the second-highest number of comments (26 with eight commenters). This article was also authored by Blatt and published in *Plant Physiology*. PubPeer received 371 comments about the article. Blatt—an unregistered commenter on PubPeer—cross-posted his PubMed Comment on PubPeer. His comment was identified as an author response on PubPeer (see Figure 4.13). This practice of identifying author responses was not used on PubMed Commons.

Figure 4.13

PubPeer Post Identifying Blatt's Comment as an Author Response



#149 Unregistered Submission

Author Response

Dear all,
I am cross-posting my response to Dr. da Silva (below) as a courtesy. You will find this post along with Dr. da Silva's letter on PubMed Commons where I welcome discussion and will engage. One other note, as I received a (pseudonymous) request by email for correction of the typo in Sarkar's name, the amendment should be online shortly.
Mike Blatt

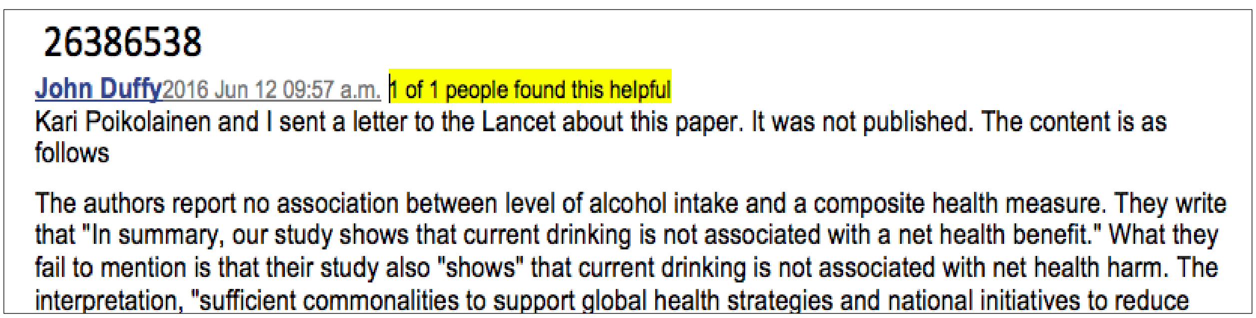
Dear Dr. da Silva,

Thank you for your comments. I am happy to discuss thoughtfully on the open forum of PubMed Commons. Like you, I reflected long and hard before writing my piece for precisely the reasons we are entering into this discussion: anonymity, fear, and scientific debate. Several of my closest colleagues expressed their anxieties that putting my head 'above the social media parapet' could have negative consequences for them. I, too, had my doubts but felt it important to raise concerns expressed to me by colleagues, young and old, which I share. I am therefore writing both in response to your invitation and generally to some of the comments posted to my editorial. I want to thank Philip Moriarty who I have come to know over the past week and who is much more eloquent than I could ever be, and Leonid Schneider who has shown true grace in stepping back from his initial cynicism to participate in the discussion.

As previously mentioned, another indicator of commenter engagement was the number of people who found a comment helpful. This feature was highlighted in a November 26, 2013, PubMed Commons Blog post, which explained that helpfulness ratings influenced the comment stream on the blog's homepage. Ratings were later used to influence article/comment placement on the PubMed GUI interface. A November 1, 2013, blog post on NCBI Insights announced that "there are now enough data in the system for people's ratings to begin having an influence on the commenting stream on the home page" (para 9). Figure 4.14 shows a screenshot of a comment that one person found helpful. Users who indicated helpfulness could remain anonymous, which somewhat diluted public awareness of PubMed Commons users and user engagement (i.e., users could not be identified, thus not counted).

Figure 4.14

Helpful Rating System on the PubMed GUI



26386538

John Duffy 2016 Jun 12 09:57 a.m. 1 of 1 people found this helpful

Kari Poikolainen and I sent a letter to the Lancet about this paper. It was not published. The content is as follows

The authors report no association between level of alcohol intake and a composite health measure. They write that "In summary, our study shows that current drinking is not associated with a net health benefit." What they fail to mention is that their study also "shows" that current drinking is not associated with net health harm. The interpretation, "sufficient commonalities to support global health strategies and national initiatives to reduce

NCBI (2013) claimed that "PubMed Commons is an un-moderated commenting system, although concerns can be reported" (para. 8). However, comments "removed by moderator" were indicated on the PubMed GUI (see Figure 4.15). Moderator involvement was a factor in how PubMed Commons functioned as an online forum for scientific discussion. Rows on the commons_archive.csv file with no text suggest the comment was either deleted by the author or removed by the moderator. Likewise, some rows merely indicate that the article URL was mentioned in another comment, leaving the comment field is blank.

Figure 4.15

Evidence of Comment Editing and Removal

Jacob H. Hanna 2015 Jan 17 10:48 a.m. edited
Comment deleted by user.

Paul Bertone 2015 Jan 16 4:00 p.m. 2 of 3 people found this helpful
A critique of this paper is available on bioRxiv: dx.doi.org/10.1101/013904

PermalinkShare

-
  - **Jacob H. Hanna** 2015 Jan 17 11:06 a.m. edited

    Comment removed by moderators.

  - **Jacob H. Hanna** 2015 Jan 17 6:49 p.m. edited 0 of 2 people found this helpful

    1) Hanna group response can be found on bioRxiv: http://dx.doi.org/10.1101/013961 It should be noted that the above speculative critique by Bertone & co. has been rejected by Nature editors and reviewers following, among other things, our response indicated above.

According to PubMed Commons editor Bastian (personal communication, October 28, 2016), comments on the forum were automatically filtered for offensive language, and then immediately posted to the PubMed GUI (i.e., not moderated before they were visible to PubMed users). Bastian and co-editor Vaughn served as moderators and manually reviewed each comment, removing those that violated forum guidelines in an effort to avoid commenter risk of being sued. If there was any question about removing a comment, NCBI director David Lipman made the final decision. Either Bastian or Vaughn was required to explain why removing a comment was actionable and if they could not make a case, the comment was not removed from the PubMed Commons forum. Given that a comment was found actionable, the commenter was contacted by the moderators and offered the choice to pull or edit their comment. Bastian disclosed that most commenters were willing to edit their comment because “they care and really want their information out there” (H. Bastian, personal communication, October 28, 2016). She added that most commenters were not intentionally mean or rude.

Bastian finished her insights about moderating comments saying, “some might disagree with the threshold of moderating [on PubMed Commons]” (H. Bastian, personal communication, October 28, 2016).

Bastian further explained that PubMed Commons guidelines were “complex.” She revealed that the word “inflammatory” had been added soon after the forum was introduced. She shared that she and Vaughn communicated to commenters that,

> NCBI is looking out for you and protecting you from potential consequences. The objective of site guidelines was to raise the bar for politeness. Because of the tendency for aggressive communication, PubMed Commons leaders wanted the forum to be a safe place. I was hesitant at first about PubMed Commons. Others were not. Authors are thin-skinned. (H. Bastian, personal communication, October 28, 2016)

Showing which comments were deleted by the commenter or removed by the moderator gave PubMed users and PubMed Commons commenters a good idea of how comments were evaluated. Accordingly, users could better judge the ability to engage in uncensored discussions on the forum.

It is difficult to determine the number of posts with moderator involvement, as commenters were first invited to edit comments that violated forum guidelines. The PubMed GUI indicated only that a comment had been edited, not that the author did it of their own volition or if he/she was invited to do so. Knowing this information, the commenting thread shown above in Figure 4.15 might indicate that Jacob Hanna made an initial comment he was invited to edit. The edited comment did not adhere to commenting guidelines and was removed by moderators. Hanna’s response post to Paul Bertone was also edited and then removed by moderators. Hanna was persistent and finally posted an acceptable comment.

Table 4.10 shows that of the 2,441 threaded comments (i.e., articles that received two or more comments), 34% (838) were edited by the commenter; 2.4% (60) were publicly removed by moderators; and 1.7% (41) were deleted by the commenter. Another critical point about moderating on PubMed Commons is that readers were unable to see if a person's comment(s) were immediately blocked based on past commenting behavior or if a commenter had lost their posting privileges (H. Bastian, personal communication, October 28, 2016). Since this information was not documented anywhere, it is impossible to report the effect this factor might have had on PubMed Commons commenting.

Table 4.10

Number of Threaded Comments Edited, Deleted, or Removed

Comments	Edited by User	Deleted by User	Removed by Moderator
5+	175	16	14
4	66	4	3
3	357	40	24
2	241	21	10
Total	838	81	60

Several social media posts I discovered in my systematized search for information about PubMed Commons indicated that commenters had indeed been contacted about editing their posts and/or were blocked from commenting altogether (see Figure 4.16).

Figure 4.17 shows that when calculating publicly available data for moderator involvement for all comments on PubMed Commons, including single comments on a PubMed indexed article, more comments were deleted by a commenter than were removed by moderators.

Figure 4.16

Social Media Posts Reporting Moderator Involvement

#22 John A. Greenwood Author Response

Dear Lydia,

My condolences on your frozen PubMed account. I agree that the selection of comments that were deleted is odd - they seem no different to the others to me. That said, I don't see why this is the appropriate place to blog about that.

Sin Hang Lee, MD
February 9, 2018 at 7:41 pm
Reply

Just for the record. After I posted this comment to several Abstracts under which I posted comments previously, my PubMed Commons account is suspended.

The travesties that are Glantz, epidemiology modeling, and PubMed Commons

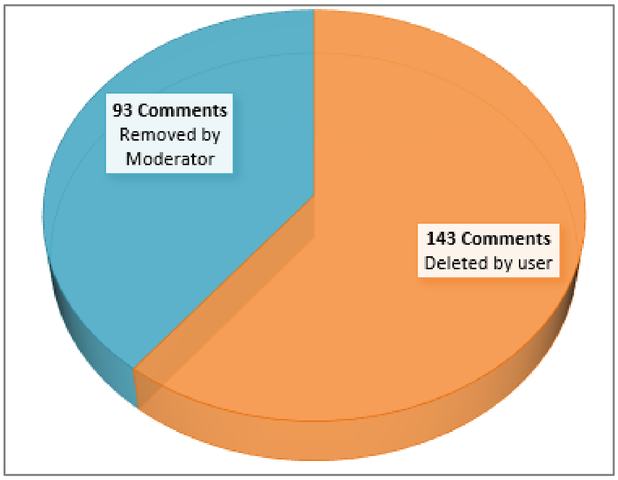
Posted on 6 January 2018 by Carl V Phillips | 12 Comments

by Carl V Phillips

On top of everything else, the single message I got from PMC contained a rather rude threat to pull my credentials to comment there if I again violated those vague rules they refused to clarify. That would create a rather difficult situation for someone who wanted to continue to give them free content. Fortunately for me, I am not inclined to gift them any more.

Figure 4.17

Publicly Available Number of Comments Deleted or Removed



The next section builds upon data reported in the comment-level analysis of PubMed Commons and delves deeper into the demographics and posting characteristics of forum commenters.

4.6 RQ1b: Characteristics of Individuals Commenting on PubMed Commons

Forum overview-, journal-, and comment-level data analyses signaled that commenter demographics, characteristics, and posting behavior might have affected how PubMed Commons functioned as an online forum and the decision to discontinue it.

Combining information captured on the PubMed GUI with data from CSV files drew attention to usage patterns and commenter characteristics. Data collection and analysis from Burns and Farabough in 2015 spawned nagging questions that prompted data collection about commenters after the forum ceased to operate in February 2018. For example, were commenters' publication history or posting time of day antecedents to commenting behavior or content of the comment? Which countries and organizations were commenters associated with? Were commenters more likely to be male or female? Did commenters typically have less or more professional experience? Observations and notes from previous analyses informed a list of 25 descriptors to build a relational database in Dedoose.

4.6.1 Dedoose

To better analyze factors affecting PubMed Commons, I selected a statistically significant sample of 381 of the 7,629 comments (5%) that comprised PubMed Commons. The sample size ensured a 95% confidence level with a 3% margin of error. I entered data related to various characteristics associated with commenters and their comments, building a database on Dedoose (https://www.dedoose.com/).

Dedoose is a web-based Software as a Service (SaaS) application for mixed methods research developed by academics from the University of California in Los Angeles (UCLA) with support from the William T. Grant Foundation. In addition to traditional qualitative data management and excerpting/coding, the platform offers surface pattern analyses showing information not easily seen otherwise by integrating demographics, field descriptors, and code weighting. Over the years that I have investigated PubMed Commons, I have continued using Dedoose as it evolved. Final data analysis was performed using Dedoose 9.0.

An October 28, 2016, meeting with PubMed Commons co-editors Bastian and Vaught revealed that the two had also been thinking about data that should be collected about the forum. Vaught mentioned that she and Bastian had noticed a “lull” in the number of comments on holidays, although she said that posts on those days were the “worst—in other words inflammatory or rude” (M. Vaught, personal communication, October 28, 2016). Bastian added in jest that maybe the commenters had “too much time or too much alcohol” on those occasions (H. Bastian, personal communication, Oct. 28, 2016). These observations raised my concerns about the possible mitigating factor of time availability for busy scientists, clinicians, and researchers to compose comments and post them on the forum (Tennant, Apr. 12, 2017). Along these lines, I began to ponder what purpose or reward motivated commenting behavior.

Vaught shared that most comments were posted on biomedical, not medical, articles, and that comments tended to be clinical. Bastian recalled that the last time she and Vaught collected data about the forum was during a three-month period in 2014 when they manually tabulated the amount of time between posts, as well as the number of comments made per commenter who was “listed on the account” (H. Bastian, personal communication, October 28, 2016). Data and analyses thereof were reported during a poster presentation on Feb. 11, 2016, at the annual American Association for Advancement of Science (AAAS) (see Appendix G). Bastian added that when reviewing requests to become a PubMed Commons commenter, some individual’s credentials showed the applicant was mid-level in their career and others “weren’t academic at all” (H. Bastian, personal communication, October 28, 2016). At the time, NCBI did not track IP addresses to capture the geographic location of commenters.

Bastian (2016) admitted that data about PubMed Commons should be collected and analyzed—whether comments focused on “criticism or curation or if they were non-specific”, “type of publication an individual is commenting on” (personal communication, October 28, 2016.

4.6.2 Commenter Descriptors Associated with Comment Content

To characterize PubMed Commons commenters, data from sources listed in Table 4.11 were input into the Dedoose descriptor fields for each commenter that was part of the sample set. A student assistant helped me with data entry during 2018 and 2019. Comments listed in the “Content” column on my “commons_archive COMPLETE” Excel spreadsheet created from the commons_archive.csv file were saved into a new “comments ONLY_archive COMPLETE” Excel spreadsheet. Also, PubMed Commons comments from my “PubMed Commons pubs with comments” MS Word file were separated into the following MS Word documents: “Article with 2 comments,” “Articles with 3 comments,” “Articles with 4 comments,” and “Article with 5+ comments” (see Figure 4.18). Recall that an MS Word document was generated by copying/pasting the article abstract and comments that appeared on the PubMed GUI following my PubMed search for “all[sb]”. For this particular data analyses, I reviewed each PMID and the associated threaded/nested comments so that descriptor 7–”comment type [self reply, single, threaded, nested] and descriptor 17–”number of commenter’s comments on PubMed Commons” options could be entered into the Dedoose database.

Table 4.11

Dedoose Database Descriptors for Commenters/Comments

Desc	Set Fields	Type	Option list	Data Source
1	Line number on Excel spreadsheet	Number	N/A	commons_archive.csv Excel spreadsheet
2	Comment ID on Excel spreadsheet	Number	N/A	commons_archive.csv Excel spreadsheet
3	PMID (i.e., PubMed identification number) of article receiving comment	Number	N/A	commons_archive.csv Excel spreadsheet
4	Comment date	Date/Time	N/A	commons_archive.csv Excel spreadsheet
5	Comment time	Option list	Each hour	PubMed “all[sb]” search csv Excel spreadsheet

(table continues)

Desc	Set Fields	Type	Option list	Data Source
6	Commenter first and last name	Text	N/A	commons_archive.csv Excel spreadsheet
7	Comment type	Option list	Self-reply Single Threaded Nested	MS Word document populated with PubMed “all[sb]” search for PubMed Commons comments
8	Commenter was author of article receiving comment	Y/N	Y/N	MS Word document populated with PubMed “all[sb]” search for PubMed Commons comments
9	Commenter place of employment	Text	N/A	PubMed search Google search ORCID search
10	Commenter experience level	Option list	Early Mid Late	PubMed search Google search ORCID search
11	Commenter country of work affiliation	Text	N/A	PubMed search Google search ORCID search
12	Comment gender	Option list	Male Female Unknown	commons_archive.csv Excel spreadsheet with additional Google search
13	Commenter field of expertise	Option list	See list below*	PubMed search and Web of Science search
14	Number of PubMed-indexed publications authored by Commenter	Number	N/A	PubMed search
15	Commenter is first author on PubMed-indexed publication	Y/N	Y/N	PubMed search
16	Commenter published in same journal of publication receiving PubMed Commons comment	Y/N	Y/N	PubMed search
17	Number of commenter’s comments on PubMed Commons	Number	N/A	PubMed “all[sb]” search csv Excel spreadsheet sorted by commenter name
18	Number of Commenter’s comments on PubPeer	Number	N/A	PubPeer Google Chrome Extension search and PubPeer website search

(table continues)

Desc	Set Fields	Type	Option list	Data Source
19	Title of journal in which publication receiving PubMed Commons comment was published	Text	N/A	PubMed "all[sb]" search csv Excel spreadsheet sorted by commenter name
20	Field of journal in which publication receiving PubMed Commons comment was published	Option list	Same as Descriptor 13	PubsHub Journals and Congresses database search
21	Title of article receiving a comment	Text	N/A	PubMed "all[sb]" search csv Excel spreadsheet sorted by commenter name
22	Publication year of publication receiving PubMed Commons comment	Number	N/A	PubMed "all[sb]" search csv Excel spreadsheet sorted by commenter name
23	Open access publication receiving PubMed Commons comment	Y/N	Y/N	PubMed "all[sb]" search csv Excel spreadsheet sorted by commenter name
24	Altmetric Attention score for publication receiving PubMed Commons comment	Number	N/A	Dimensions Digital Science and Research Solutions, Inc. database search
25	Dimensions score for publication receiving PubMed Commons comment	Number	N/A	Dimensions Digital Science and Research Solutions, Inc. database search

*Option list for descriptors 7 and 17: Acupuncture, Addiction–Pain, Administration, Aging, Allergy, Anesthesiology, Behavior–Psych, Biology, Biotech, CAM Therapy, Cardiology, Chemistry, Dentistry, Dermatology, Emergency Medicine, Environ. Health, Gastroenterology, General Medicine, Genomics, Health Services, Hematology, Immunology, Info Science, Medical Education, Medical Tech, Neurology, Nursing, Nutrition-Obesity, OB/GYN, Oncology, Ophthalmology, Orthopedics, Osteo, Other, Otolaryngology/ENT, Pathology, Pediatrics, Pharma, Podiatry, Radiology, Research, Sleep Medicine, Social Work, Sports Medicine, Surgery, Toxicology, Transplantation, Urology, Vet Medicine

Figure 4.18

MS Word Document with Data Entered into Dedoose

Articles with 2 comments

25286440

KLAUS KAESTNER2015 Mar 27 5:49 p.m. 3 of 3 people found this helpful
Dear Dr. Tarlow We are well aware of the fact that a CreER is required for genetic lineage tracing. We tried many times to derive a Foxl1-CreER line. Unfortunately, none of our six founders showed any expression. We will keep trying!

In the case of Foxl1 expression in the liver, there is an additional difficulty to consider. There are no Foxl1 expressing cell in the healthy, uninjured liver. Thus, even a Foxl1-CreER does not help, as it will not label any cell! The Foxl1 promoter becomes activated only AFTER specific types of liver injury.

Note that additional evidence for the bi-lineage potential of Foxl1-Cre marked cells comes from the fact that one can establish clones of cells from a SINGLE Foxl1-Cre/RosaYFP labelled cell, and these can be expanded indefinitely. Thus, these cells are clonogenic. In vitro, these cells can be differentiated towards both the cholangiocyte and hepatocyte lineage.

Note also that we acknowledge the limitations of our current model in the discussion of this paper. Klaus Kaestner

PermalinkShare

Branden David Tarlow2014 Dec 22 11:10 p.m. 1 of 1 people found this helpful
It's unclear why this group has not replicated their results with the Foxl1-Cre with a CreERT2 allele since the original 2009 publication. I agree with a recent review that stated "Experiments using non-inducible Cre lines do not constitute bona fide lineage tracing tools..." (see Lemaigre Hepatology 2014; doi: 10.1002/hep.27659)

PermalinkShare

28526097

James M Heilman2017 Jul 29 5:47 p.m. 3 of 3 people found this helpful
Seriously "sea-buckthorn oil protects against infections, prevents allergies, eliminates inflammation and inhibits the aging process". This sounds like world changing news. Looking for the RCTs that back it up and not finding any. The only RCT listed in the refs found NO benefit. https://www.ncbi.nlm.nih.gov/pubmed/23131570 That paper which found NO benefit is used to support this sentence "Sea-buckthorn oil as well as extracts from its fruit are used as an adjunctive therapy in treatment of many diseases". If it has no effect that is not a treatment.

PermalinkShare

- **Stuart RAY**2017 Aug 14 10:10 a.m.
  These are substantial concerns - I have written to Professor Nowak to invite comment on this discussion.

  PermalinkShare

After data for the 25 descriptor fields were entered for each of the 381 commenters in the dataset, corresponding comments from the "comments ONLY_archive COMPLETE" Excel spreadsheet (described above) were uploaded as "media" into Dedoose. This enabled each comment to be associated with the individual who posted the comment so that descriptor fields could be analyzed on both a collective and granular basis.

Automated database processing verified my manually tabulated findings. Pie charts and graphs in the figures below visualize the findings, some which reiterate findings from the quantitative content analysis reported in the previous subsection. Pie charts in Figures 4.19 and 4.20 show that most commenters were the first author of an article that was indexed in PubMed, although they were not the author of the article he/she commented on. Less than half of the commenters commented on articles published in a journal that he/she had also published in. This raised concerns about “peer review,” especially since the only requirement for being a commenter on PubMed Commons was being an author on an article indexed in PubMed. Chapter 5 looks at this muddy concept of “peer” as a mitigating factor for forum use. Commenters were chiefly associated with an organization or university located in the United States, followed by the UK, Canada, and France. Most commenters were in the mature (i.e., termed “full” in the database) phase of tenure in their career, and most articles that received comments were open access (i.e., freely available without the need for a subscription to the journal that published the article).

As previously reported, the majority of commenters were associated with a single comment, and most PubMed-indexed articles received two comments.

Data collected for the sample set analyzed in Dedoose also confirmed the quantitative content analysis of the entire forum, showing that commenting behavior was relatively steady from May 2014 to Mar. 2017 (see Figure 4.21).

Figure 4.19

Dedoose Field Data Visualization of Commenter Characteristics

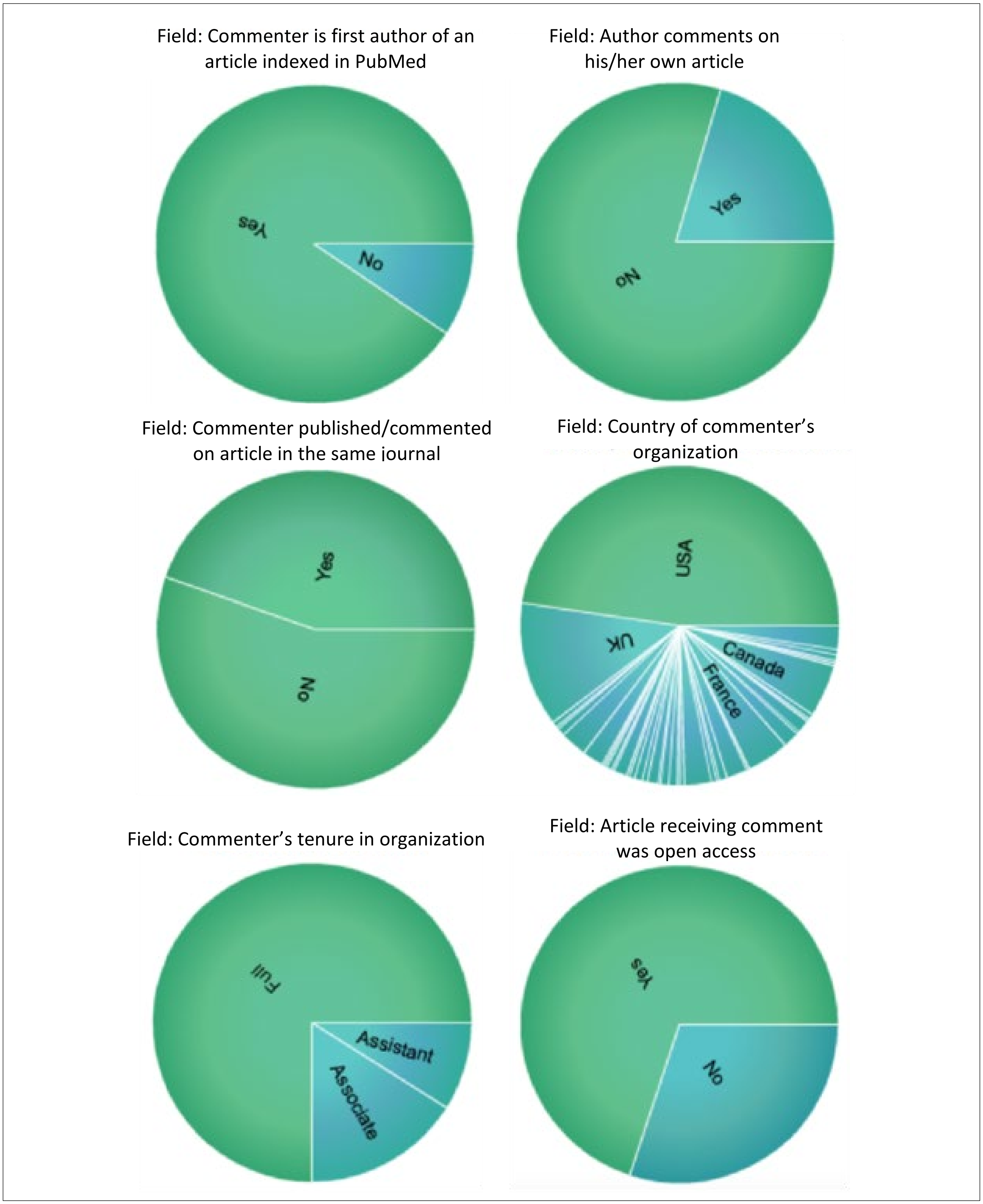


Figure 4.20

Dedoose Field Data Visualization for Number of Article Comments

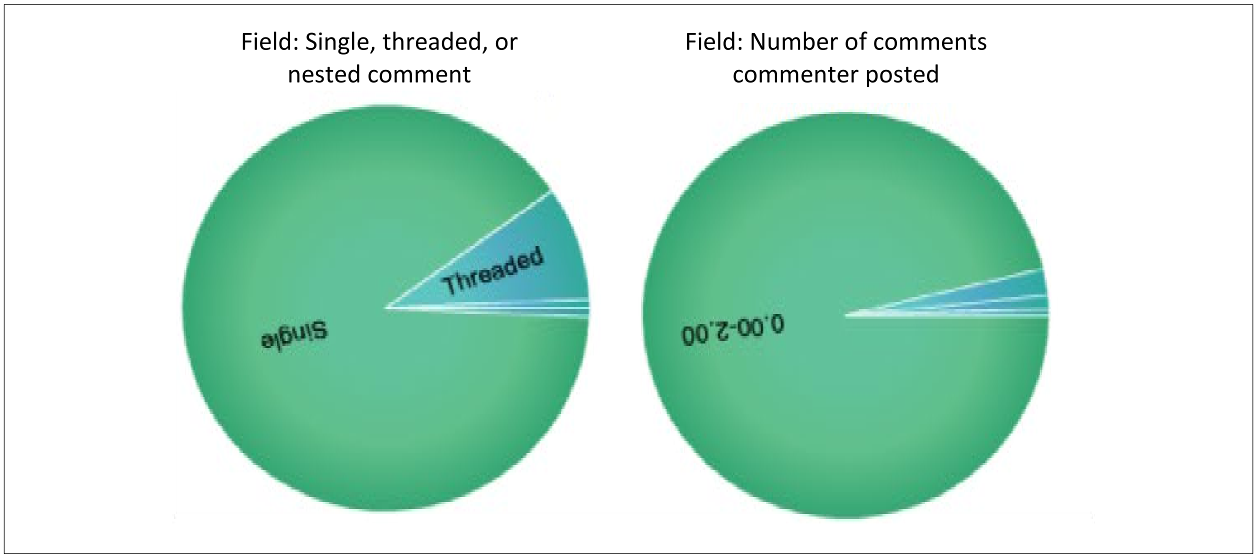
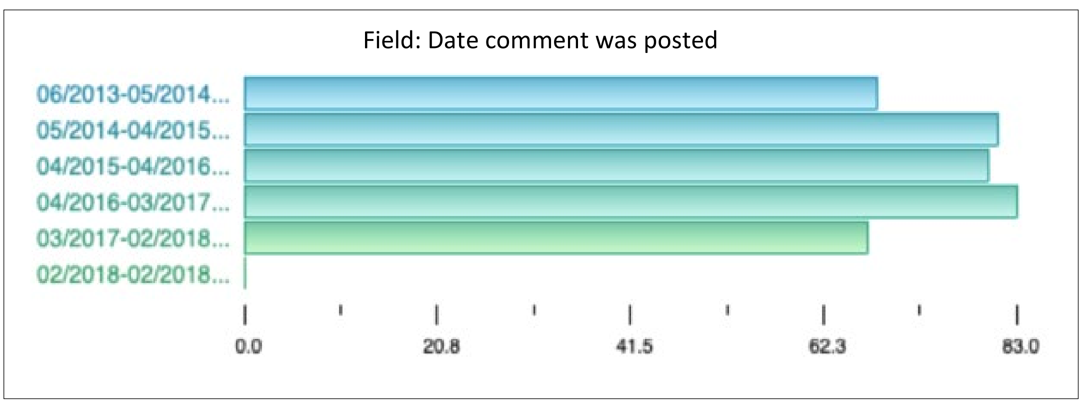


Figure 4.21

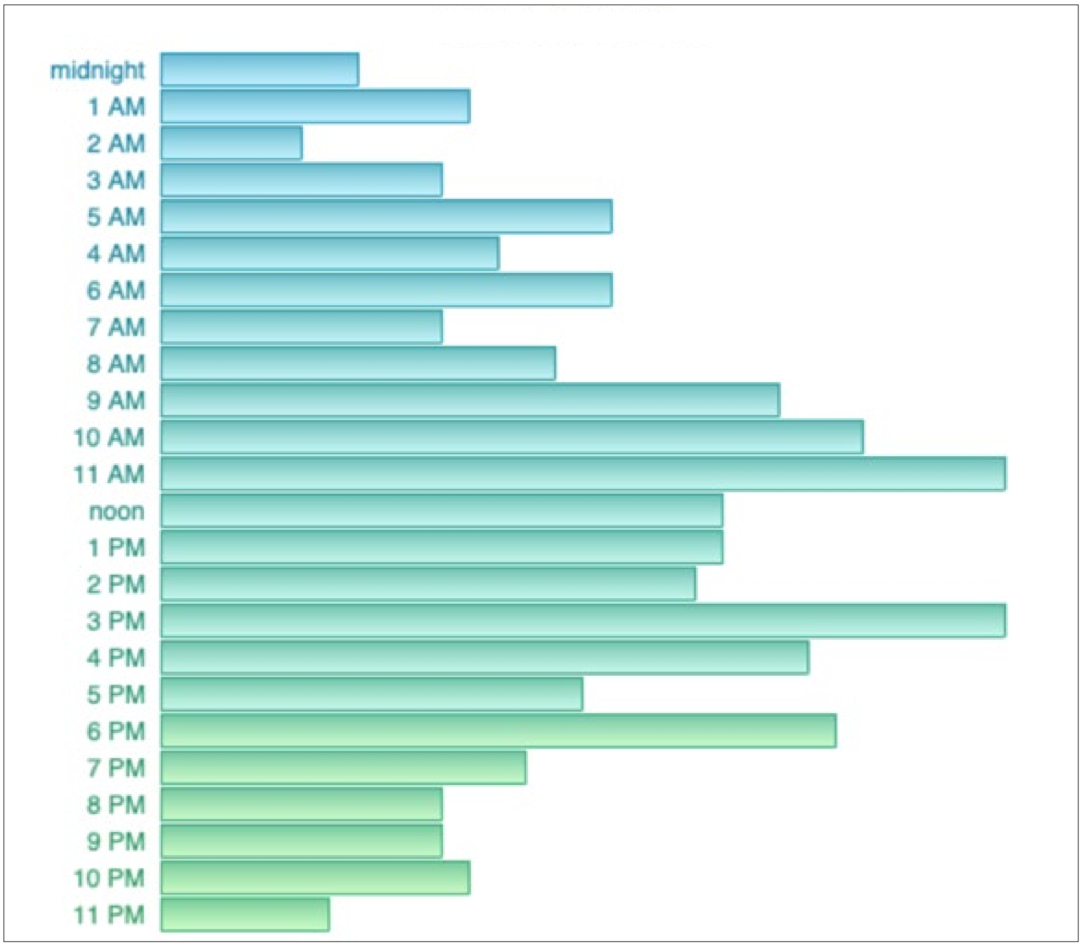
Dedoose Results for "Date Comment was Posted" Field



Commenters generally posted comments during the daytime hours, with 11 a.m., 3 p.m., 6 p.m., and 10 a.m. (in descending order) being popular posting times (see Figure 4.22). Overwhelmingly, commenters were not the authors of the articles they were commenting on, and the articles they commented on were open access (i.e., available to readers at no cost). As previously indicated, articles typically received a single comment, and commenters posted an average of two comments (see Figure 4.20).

Figure 4.22

Dedoose Results for "Time-of-Day Comment was Posted" Field

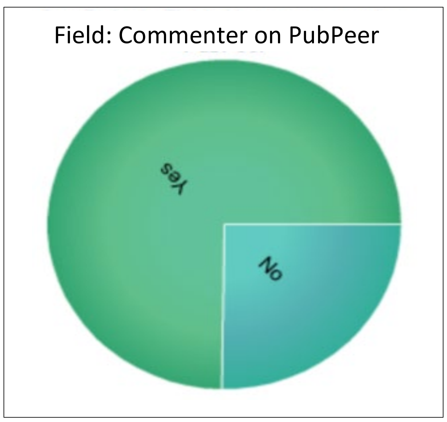


A comparison of comments posted on PubPeer revealed that most, but not all, PubMed Commons' comments were also visible on the PubPeer website (see Figure 4.23). Nearly 75% of PubMed Commons commenters were identified as commenters on PubPeer. The high incidence of cross posting is likely to do automated downloading of PubMed Comments onto PubPeer. Alternative online forums for scientific discussion and PPPR, like PubPeer, are discussed in Chapter 5.

Even more important than validating the manual counting reported from my quantitative content analysis, the automated analysis of the PubMed Commons sample dataset in Dedoose revealed insightful associations among commenter characteristics. The next section shares results from linking comment content with commenter demographics and characteristics.

Figure 4.23

Dedoose Results for "Commenter Also Appears on PubPeer" Field

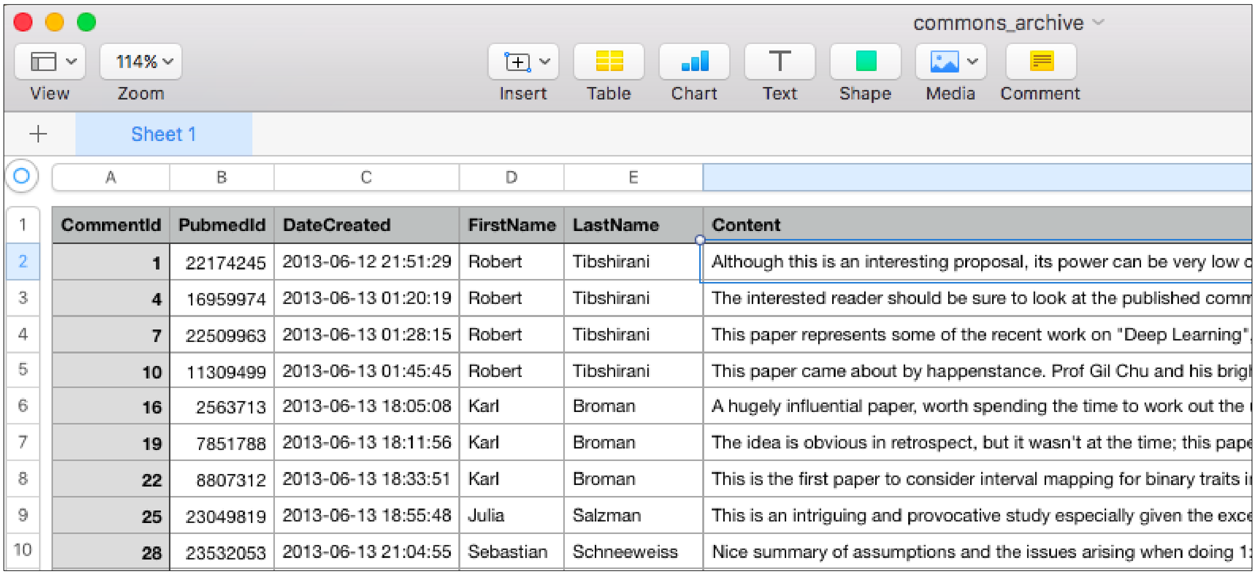


4.7 RQ1c: Subject Matter of Comments Posted on PubMed Commons

Comments from PubMed Commons were downloaded as a CSV file from the NCBI FTP site on March 3, 2018—the last day the forum was visible on the PubMed GUI (see Figure 4.24). The "commons_archive.csv" file was composed of 7,629 comments that were posted on the 6,013 PubMed-indexed articles between June 12, 2013, and February 18, 2018. Note that PubMed Commons innovator Rob Tibshirani made the first three comments on the forum. Interested readers can still access the entire raw dataset of commenters and their associated comments from the link to the NCBI FTP site found at https://www.ncbi.nlm.nih.gov/pubmedcommons/.

Figure 4.24

Download of commons_archive.csv File



	A	B	C	D	E	
1	CommentId	PubmedId	DateCreated	FirstName	LastName	Content
2	1	22174245	2013-06-12 21:51:29	Robert	Tibshirani	Although this is an interesting proposal, its power can be very low c
3	4	16959974	2013-06-13 01:20:19	Robert	Tibshirani	The interested reader should be sure to look at the published comm
4	7	22509963	2013-06-13 01:28:15	Robert	Tibshirani	This paper represents some of the recent work on "Deep Learning".
5	10	11309499	2013-06-13 01:45:45	Robert	Tibshirani	This paper came about by happenstance. Prof Gil Chu and his brigl
6	16	2563713	2013-06-13 18:05:08	Karl	Broman	A hugely influential paper, worth spending the time to work out the
7	19	7851788	2013-06-13 18:11:56	Karl	Broman	The idea is obvious in retrospect, but it wasn't at the time; this pape
8	22	8807312	2013-06-13 18:33:51	Karl	Broman	This is the first paper to consider interval mapping for binary traits i
9	25	23049819	2013-06-13 18:55:48	Julia	Salzman	This is an intriguing and provocative study especially given the exce
10	28	23532053	2013-06-13 21:04:55	Sebastian	Schneeweiss	Nice summary of assumptions and the issues arising when doing 1:

Rows of text in the comment “Content” column on the PubMed Commons CSV file were saved in a separate MS Excel spreadsheet (N=7,629), and then uploaded as a media file into the “PubMed Commons forum” project database I had created using an earlier version (8.2.14) of Dedoose. Every twentieth comment, beginning with Commentid 4 (i.e., the second PubMed Commons comment written June 12, 2013, at 1:20 a.m. by Robert Tibshirani) was associated with the commenter’s name, and 25 descriptor fields were subsequently populated for each commenter.

The primary goal for this particular data analysis was capturing comment sentiment. To a lesser degree, I was also interested in discovering meaningful insights about commenters and how certain characteristics affected the nature (i.e., content) of their comments.

Code Development. Comment codes were iteratively developed beginning with a 2015 pilot study of PubMed Commons. My colleague Burns (SB) and I (MF) performed a content analysis of threaded comments posted on PubMed Commons, using a constant comparative method in accordance with grounded theory procedures (Glaser & Strauss, 1967). After an initial comparative analysis of 232 randomly selected comments, a list of comment topics emerged (i.e., open coding). We (SB and MF) each critically reviewed the comments and coded them line-by-line. Comments were identified, named, categorized, and described. Themes were proposed, and coding syntaxes were jointly reviewed until a saturation of mutually agreed-upon categories was observed (i.e., focused categories). Axial coding was adopted to further analyze comments and discover interrelationships among codes. Some properties and dimensions of certain categories were linked, and subcategories were created for others. For example, “disagreeing” was further broken down into “disputing” and “flaming” to capture the intensity of disagreement. An “inquiry” subcategory was added to identify instances when the commenter was looking for clarification from article authors, and a “redirecting” subcategory was added to indicate when a commenter added a link or provided an article reference in their comment. Minor discrepancies were resolved through discussion and consultation until 11 mutually agreed-upon comment codes were determined: public forum, discussing, disputing, redirecting, humor, watchdog, flaming, validating, inquiring, promoting, and author responding.

In 2016 at the annual American Association for Advancement of Science (AAAS) National Conference, PubMed Commons co-editors Bastian and Vaught (2016) presented a poster that identified seven comment codes that described forum posting content: endorsement, curation, critique, discussion, author addendum, author reply, and author update/revision (see Appendix G). I considered these categories and definitions as I finalized my decision to code comment content into 12 categories (see Table 4.12). Iterative data collection and analysis, as well as the ongoing systematic search for literature about the forum, urged me to further divide the "discussion" category into "post-publication peer review" and "alternative for scholarly communication." For the former, there had been an intensifying focus and heated discussions by PubMed Commons commenters who seemingly interpreted the sole purpose of the forum was post-publication peer review. This perception had been further perpetuated by PubMed Commons organizers, who introduced and began regularly using the term "post-publication" in tandem with PubMed Commons (Allison, Brown, George, & Kaiser, 2016; Hasan, Masood, & Mumtaz, 2016; Lane, 2016; NCBI Insights, 2017; PubMed Commons Team, 2017; Teixeira da Silva, Al-Khatib, & Dobranszki, 2017; NLM Technical Bulletin, 2017). With regard to creating a category for "alternative for scholarly communication," it became obvious that the forum was used a vehicle for making unpublished letters to the editor publicly available. A review of comments showed that many posts had citations and a formatted list of references, which are marks of scholarly communication.

Table 4.12

Development and Definition of Qualitative Content Analysis Categories

Farabough & Burns (2015)	Vaughn & Bastian (2016)	Farabough Dissertation (2019)	Definition
Author responding	Author addendum	Author reply	Author provides a correction, replication, or revised interpretation to the article commented on
Author responding	Author reply	Author update/revision	Response/reply by the article author—directed toward another commenter—addressing point(s) raised in a previous PubMed Commons comment

(table continues)

Farabough & Burns (2015)	Vaughn & Bastian (2016)	Farabough Dissertation (2019)	Definition
Author responding	Author update	Author addendum	Author adds further information about the article commented on (e.g., links to supplemental information, media coverage)
Disputing	Critique	Critique	Discussion that includes an argument about the article, perhaps with unfounded personal attacks
Redirecting	Curation	Curation	Annotation of information, resources, literature relevant to evaluation, interpretation, reproducibility of the article (i.e., organization and integration of facets of the article)
Discussing/ Public Forum	Discussion	Discussion	Commenting without criticism of the article
Validating/ Promoting	Endorsement	Endorsement	Short, favorable notation of the article
Humor	—	Humor	Commenter purposefully initiates levity/humor in the discourse of the comment about the article
Inquiring	—	Inquiry	Direct question seeking a response from author or another commenter about the article
Discussing	Discussion	PPPR	Fully constructed review of the article that evaluates the article or delivers feedback to authors on the merits of research, methodology, argument, and/or conclusions in an effort to monitor, shape, advance the field, or lay the groundwork for accepting new/novel ideas
Watchdog	—	Watchdog	Reporting a correction, retraction, erratum, or other official scholarly communication notice about the article
Discussing	Discussion	Alternative to formal scholarly communication	Substitute for letter to the editor or separate article on topic of the article

A clearer definition of comment content based on data collection and analyses made for effective evaluation of the PubMed Commons forum as a vehicle for scientific discussion. This was a critical process for identifying mitigating factors for the forum's use/disuse. Comments were coded in “message thoughts” units of analysis, as most comments were composed of several topics. For example, consider the change in thought for the following comment. “According to the information at ClinicalTrials.gov (http://clincialtrials.gov/show/NCT01024231) the primary endpoint of the trial was safety (as expected for a phase 1 trial), and this primary endpoint is mentioned neither in the abstract nor the methods section.” This first part was coded as curation, and the part following the comma was coded as critique.

4.7.1 Interrater Reliability

Valid and reliable research studies must include procedures that measure agreement among researchers who are collecting and/or analyzing data. Consistency is of concern due to the variability among humans and their thought processes. Perfect agreement is infrequently achieved. Confidence in study results is a consequence of the amount of disagreement in tandem with the error introduced from inconsistency in coding assigned by coders. The extent of agreement among data analysts is commonly referred to as interrater reliability.

To validate my codes, I randomly selected 40 comments from among those not included in the sample set of 381. The 40 comments were imported into an Excel spreadsheet with comments organized in rows and the 12 categories organized in columns. Burns and I coded the comments independently on identical Excel spreadsheets, and then met virtually to resolve conflicts. Interrater agreement on the test set was 84%. Adjustments were made to category definitions, and a revised codebook with code terms/definitions was used for coder training.

The 12 coding categories defined in Table 4.12 were then entered as another set of descriptors in the Dedoose database, where I read through the sample set and coded message thoughts. The 381-comment sample set was also saved on an Excel spreadsheet in the same format as the one described above, and then emailed to JC, who randomly selected and coded 40/381 comments (approximately 10% of the full sample set). Interrater agreement on this dataset was 97.5%.

Table 4.13 categorizes comment message thoughts (n=782) from the 381 comments that were analyzed. Percentages have been adjusted based on the total number of message thoughts coded from the 381 comments (n=782).

Table 4.13

Distribution of Sample Set Message Thought Categories

Code	Message Thoughts	Percent of All Message Thoughts (%)
Author reply	29	3.7
Author update/revision	9	1.1
Author addendum	33	4.2
Critique	161	20.6
Curation	203	25.9
Discussion	150	19.2
Endorsement	50	6.4
Humor	1	<1
Inquiry	30	3.8
PPPR	60	7.7
Watchdog	21	2.9
Alternative to formal scholarly communication	35	4.5
TOTAL	782	100%

4.7.2 Examples of Comments

Following are examples of each comment category. For ease of reading, the term for each code is followed by the definition listed in Table 4.12.

- Author Addendum: author provides a correction, replication, or revised interpretation to the article commented on. For example:

> We have a new version of Figure 6 that fixes a mistake in two of the images that have the wrong chirality due to the image being flipped. The corrected image can be found on a FigShare repo for the paper:
> https://figshare.com/articles/Zn_metalloprotein_paper/4229333
>
> The URLs in this paper for software availability and data availability are out of date; here are the updated URLs: The ShortStack program is now at github... the latest release is at https://github.com/MikeAxtell/ShortStack/releases
>
> The datasets used in this 2013 paper are now at https://psu.app.box.com/v/axtelldata in directory 'ShortStack_Paper_Data' A tutorial and test data for ShortStack are at https://psu.app.box.com/v/axtelldata in directory 'ShortStack_TestData.'
>
> Finally, I would like to point out that the current version of ShortStack is much enhanced relative to what was described in this 2013 paper. Many of the advancements are described by my group in Johnson et al. (2016):
> https://www.ncbi.nlm.nih.gov/pubmed/27175019.
> Thanks, Mike Axtell

- Author Reply: response/reply by the article author—directed toward another commenter—addressing point(s) raised in a previous PubMed Commons comment. For example:

> There are many interesting follow-ups to our work indeed, some of which I believe merit further study. You have begun to identify some of these, and it seems to me there is the possibility for a constructive conversation to be had here. I highly encourage you to stop by our upcoming poster at the Vision Sciences Society conference in Florida in May, where we extend this work by exploring electrophysiological correlates of performance on this task in various conditions in both age groups.
>
> Our research group would be happy to discuss the issues you are taking with our work, as well as potentially-fruitful follow-ups that can further address the questions we have raised in this work and that you have touched in some of the above comments.I believe that, especially due to the presentation of this work over a number of conferences where I was challenged by experts in the field who helped me formulate and refine the ideas presented, as well as the rigorous peer review editorial process leading to the publication of this work in a high quality journal, the rational of our hypothesis and interpretation of our results are clearly laid out in the paper.
>
> Thank you again for your keen interest in this work.

Dear Joshua, thank you again for your comments. I am worried that you continue to cut and paste to distort my sentences.

1. The headline over my text was written by the Nature editors as their introduction to the paper, so perhaps you should blame them and ask them to replace it with “Here follows a horrible paper by Ioannidis”. Yet, I think you would still be unfair to blame them, because their headline says, “most innovative and influential”, not just “most innovative”. The terms “influential”, “influence”, “major influence” pervade my paper multiple times, but you pick one sentence with “innovative” instead and interpret it entirely out of its context.

2. The phrases “the most important” and “very important” are not identical. Very important papers may not necessarily be THE most important. But they are very important - and influential. [As an aside, honestly, this repeated cross-examining quotation-comment style makes me feel as if I am answering the Spanish Inquisition. Am I going to be burnt at the stake now (please!) or there is one more round of torture?]

3. We agree we need evidence, more evidence - evidence is good, on everything, including the current NIH funding system, which has practically no evidence that it better than other options, but still distributes tens of billions of dollars per year. Wisely, I am sure.

4. “your list contains...”. This is not my list. This is the Scopus list. Right or wrong, I preferred not to manipulate it. Your colleagues did manipulate it and did not even share the data on how exactly they manipulated it.

5. You continue to use the term “innovative thinker” out of its context. I scanned again carefully my paper and I can’t find the word “excellent”. In my mind, a student who has authored as first author a paper that got over 1000 citations (and the paper is not wrong/refuted) is already worthy to be given a shot as a principal investigator. If you disagree, what can I say, feel free not to fund him/her.

And please don’t worry, most of these guys are not funded anyhow currently, many of them even quit science. Hundreds of principal investigators who publish absolutely nothing or publish nothing with any substantial impact get funded again and again. Hurray! I am afraid it is unlikely there will be more convergence in our views at this point. A million thanks once again, I have learnt a lot from your comments.
John

- Author Update/Revision: author adds further information about the article

commented on (e.g., links to supplemental information, media coverage). For example:

After reviewing the entire paper, we noticed an error in data of the last column of Table 2. During the registration of information in Table 2, the last column mistakenly recorded incorrect monthly number of deaths for myocardial infarction. The correction will be done this week. Once this is a government data, it can be found in http://tabnet.datasus.gov.br/cgi/tabcgi.exe?sim/cnv/obt10SP.def.

> Additional disclosure. One of the co-authors, R.A., has an equity interest in Molsoft, LLC. The terms of this arrangement have been reviewed and approved by the University of California, San Diego in accordance with its conflict-of-interest policies.

- Critique: discussion that includes an argument about the article. For example:

> I believe the claims regarding “beneficial” organisms in this paper are inaccurate and misleading. For example, consider the statement in the abstract: “While the vast majority of microbial species classified were beneficial “No evidence is presented anywhere in the paper that the microbes they identify via sequence analysis are beneficial in any way.
>
> This study funded by members of the International Association of Color Manufacturers (IACM) and written by IACM staff, members, and consultants touting the safety of food dyes is so riddled with inaccuracies and misleading statements that it should be retracted and disregarded.

- Curation: annotation of information, resources, literature relevant to evaluation, interpretation, reproducibility of the article (i.e., organization and integration of facets of the article).

Recall that most comments redirected people to another information source (e.g., dataset, updated/revised results, blog, social media discussion, or branded website). Digital media curation is in some ways a new form of gate watching that involves identifying, selecting, verifying, organizing, describing, contextualizing, maintaining, and preserving existing content artifacts from third party websites and integrating them into a holistic resource by providing hyperlinks to redirect readers (Stanoevska-Slabeva, Sacco, & Giardina, 2012, p. 12). For example:

> Although this is an interesting proposal, its power can be very low compared to competing methods. Details can be found in our published comment in Science (web version), also available at http://www-stat.stanford.edu/~tibs/reshef/comment.pdf. In my opinion, a better measure of non-linear dependence is ‘distance correlation’: https://en.wikipedia.org/wiki/Distance_correlation Rob Tibshirani.

In a post to the 'Had I Been a Reviewer' blog, Lucy Cragg and I raised four questions about the statistical analyses used in this paper. Briefly, we ask about (i) the appropriateness of using one-tailed tests, (ii) the sampling methods, (iii) the appropriateness of using a Fisher's r-to-z transform to compare dependent correlation coefficients, and (iv) the lack of Bonferroni corrections. The full post is available here: http://ow.ly/qCri6.

- Discussion: commenting without criticism of the article. For example:

As a male physician, I will continue to take a MVMS, based on Statement 1, unless evidence emerges which disproves the results of these 2 large trials. While awaiting further information and considering the minimal potential harms and cost of multivitamins, and the possible benefits, I see no reason to dissuade women from taking a MVMS.

Dear Michael, There is no real contradiction on the format negotiations. After some to-and-fro, your final offers were indeed relatively generous given "journal constraints". But by that time, we had come to realise that we didn't need to satisfy ourselves with the "halfway" we were working towards. As anybody who has tried to correspond with a journal knows, the process can feel extremely restrictive compared to the freedom and immediacy of a blog post. Anyway, the point of the above comment was to correct rapidly three possible implications ambiguously left open (and predictably seized upon by a twitter denizen): i) that you'd offered to give us equal airtime, ii) spontaneously, and iii) that we hadn't felt able to counter your arguments. That's why I gave a bit more background about the process.

- Endorsement: short, favorable notation of the article. For example:

This kind of perspective from a cardiovascular scientist is most beneficial to general internists like me.
Timely and necessary reflection.

This is a very high-quality paper. The validation metrics are appropriate and statistically validated. The data collection is high quality, and the model supports the conclusions the authors assert. The incorporation of patient punctuality, and separating it by new and follow-up distributions, is a particularly elegant piece. And then, finally, they showed what happened when a simulated strategy was implemented in the real world, which is missing in a lot of the literature.

- Humor: commenter purposefully initiates levity/humor in the discourse of the

comment about the article. For example:

Is there a middle ground between Dystopia and Utopia?! May be the real, organic world, with or without the internet driven education system. Methinks, as well as the brick-and-mortar business has survived, nay thrived in an internet world, so the academic community has and will- if for no other reason, than that Education is Big Business!

- Inquiry: direct question seeking a response from author or another commenter

about the article. For example:

What are the parameters of the docking box used for the proteins? How much docking runs were performed? Where is the docking energy list? Where are the affinity constants calculated from the binding free energy values? How do they relate to the experimental values?

I have a question according to methods: You write “Sexual receptivity was induced in ovariectomized females with subcutaneous (s.c.) injections of estradiol benzoate (5 µg/0.1 ml in sesame oil; Sigma-Aldrich) given 48 h pretest, and progesterone (500 µg/0.1 ml in sesame oil) 5.5 h pretest. These parameters are known to induce a state of optimal sexual receptivity, comparable to the estrus phase (Jones et al., 2013).”What amount of drug solutions did you administer?

- PPPR: fully constructed review of the article that evaluates the article or delivers

feedback to authors on the merits of research, methodology, argument, and/or conclusions to monitor, shape, advance the field, or lay the groundwork for accepting new/novel ideas. For example:

Shu et al have conducted a fascinating analysis of miRNA sequences. It is unclear, however, how this analysis relates to dietary, circulating, and “transportable” categories of RNA that are hypothesized to be absorbed from the diet in functional form. Although the stated intention was “to heavily rely on experimental data to identify features that can differentiate secreted miRNAs from the rest,” the data in question are either unreliable (the circulating miRNA data is questionable, based on only one biological and technical measurement), incompletely described (assignment of “exosomal” status based on the two cited databases is unclear) or missing (the pivotal milk uptake experiment--unless I’m missing something). Thus, the practical validity of the sequence analysis cannot be assessed. Endogenous miRNAs are classified by Shu et al as circulating or not based on a list from <PMID:20847327>.

This preliminary publication reported only one qPCR threshold cycle measurement for each of several hundred miRNAs using only one sample of pooled plasma. Other issues, such as a lack of correlation with results of other studies and a failure to detect abundant plasma miRNAs, such as miR-16 and miR-223, were previously noted in <PMID:22048406>. Thus, the “circulating” classification made by the authors is not supported by reliable data. Perhaps the authors might wish to revisit their study with a more comprehensive ranking of plasma miRNAs supported by reliable public sequencing and microarray data. Which miRNAs are packaged into extracellular vesicles is an even more complicated question than simple presence in circulation. Since the majority of miRNAs in circulation appear to be in free protein complexes, not EVs, contaminants of EV preparations have strong potential to skew experimental results. It would be helpful if the authors could clarify how they used the EVpedia and ExoCarta databases to identify EV-packaged miRNAs and how this information (presumably including abundance ranks?) was used in the study.

Also unclear was where to obtain the sequencing data from the described milk feeding experiments. Although all data were said to be found on a university website, I could not find the sequencing data there or elsewhere. A public link to these data and further clarification of how they were used to validate the findings would be very helpful, as well as consistent with journal guidelines. Perhaps I missed this link? I would note that the evidence in support of the dietary miRNA transfer hypothesis described as “unambiguous” consists of a study by the authors.

The results of this study have not been confirmed. Alternative hypotheses (<PMID:25332488>) were omitted, as well as published evidence that contradicts the hypothesis, most strikingly a recent study (<PMID:26240150>) in which no miRNA uptake was observed from milk in miR-200c and miR-375 knockout mouse pups. In conclusion, the sequencing analysis looks quite interesting, but the underlying assumptions are debatable at best.

This paper was discussed at a Journal Club at the Cicely Saunders Institute, King’s College London, on Wednesday 7th October 2015.This study is a nice example of how a discrete choice experiment (DCE) can be used in palliative and end-of-life care to assess preferences for aspects of care. It also raises some very interesting questions about the differences in priorities and the extent to which caregivers might be able to act as a proxy for patients, an important consideration for end-of life care.

Our Journal Club discussed the work required to ensure sufficient attribute identification for a robust DCE, and wondered if the attributes decided upon in this paper sufficiently captured what is most meaningful for patients and caregivers at the end of life, i.e., there was no mention of a systematic review used to develop the attributes (see Bridges et al., 2011 for an example of DCE reporting guidelines).

We also would have liked to see a table of the probit regression output for clarity on how the willingness-to-pay was calculated, and more detail on this in the methods. Furthermore, we found it confusing that the authors state in the discussion that their sample size was too small to explore interaction effects, while it appears they recruited 70% more than their minimum acceptable sample size – some explanation would have been helpful.

Lastly, we wondered about the potential risk of bias of only including those patients who knew their diagnosis. This may limit the generalisability of the findings, even to a Singapore context. We enjoyed discussing this paper and look forward to

> more papers using DCE methodology in palliative and end-of-life care. Commentary by Melinda Smith

- Watchdog: reporting a correction, retraction, erratum, or other official scholarly communication notice about the article. Recall that commenters Goldacre and Oransky took on the mission to post these types of comments. For example:

> This article should have been retracted after an investigation by The University of Maryland found this article to contain “compromised” data (a total of 26 articles in 11 journals were affected). The journal Cancer Research was informed in August 2016, according to Retraction Watch http://retractionwatch.com/2017/04/26/university-asked-numerous-retractions-eight-months-later-three-journals-done-nothing/

> A substantial fraction of the text in this article has been copied _verbatim_ from an [earlier article](http://genomebiology.com/2010/11/3/r25) which it cites. This can be demonstrated by entering the article URLs <http://genomebiology.com/2010/11/3/r25> and <http://www.biodatamining.org/content/7/1/15> into [this online tool](http://www.copyscape.com/compare.php). Also, the first 10 or so references are identical and in the same order. Discussion at [this Twitter thread](https://twitter.com/markrobinsonca/status/519476871109804032) includes images which make the similarity very apparent.

- Alternative to Formal Scholarly Communication: substitute for a letter to the editor or separate article related to the topic of the article commented on.

These comments require a tremendous amount of detail, attention, and time to compose. Comments indicate that some researchers previously submitted them for publication, but they were rejected. Commenters seized the opportunity offered by PubMed Commons to share their thoughts in a public, academic forum. Most of these comments were extremely long. For example:

> Since the JAOA denied us the opportunity to reply to the letter in print, we are posting our response here on PubMed Commons.

> The letter posted below was rejected from JAMA Internal Medicine, eliminating the possibility for post publication correction of the limitations of peer review of this paper. Effect size in depression trial could be due to inadequacies of control treatment:

Comment on Davidson and colleagues (2013)Davidson and colleagues [1] claimed benefits for a collaborative care (CC) intervention for depression that exceed not only previous CC trials, but also effect sizes for a variety of interventions for depression. Strong claims often later prove to be exaggerated or simply false [2] and deserve special scrutiny. We should keep in mind that effect sizes observed in trials are not attributes of interventions, but of comparisons between interventions and control groups. Large effect sizes can simply represent the exceptionally poor outcomes of control groups. Davison et al.'s underspecified "routine" care could simply have been inadequate care.

The authors failed to acknowledge that patients in the control group had to pay for any depression treatment, whereas it was provided free to the patients in the intervention group. This might explain that the number of patients in the "routine" care group who received a new prescription of antidepressants increased by only two, versus ten in the intervention group. Similarly, the number of patients in the routine care group that received psychotherapy increased only by seven, versus an increase of 42 in the intervention group. The low rate of increased treatment occurred in the control group despite providers having been informed of patients' depression scores. Patients were designated as "depressed" based on a self-report questionnaire. Thus, we cannot determine the extent to which patients with heightened depressive symptoms but failing to meet formal criteria for major depression were appropriately not having treatment initiated or inappropriately treated. Overall, we cannot determine whether active treatment or the mere attention and support and awareness of treatment being available free were associated with the greater improvement in the intervention group.

Moreover, most patients identified as "depressed" in the intervention group were not in remission at follow up. Difficulty interpreting results could have been anticipated at the time of the study's design. In short, results of this trial are insufficient to encourage a more ambitious trial with the same basic design, because of a lack of demonstration that any particular elements of centralized care management account for the group differences in improvement in depression that were observed, rather than inadequacies in the care provided to the control group.

1. Davidson, KW, Bigger, JT, Burg, MM, Carney, RM, Chaplin, WF, Czajkowski, S, Dornelas, E, Duer-Hefele, J, Frasure-Smith, N, Freedland, KE; Haas, DC; Allan S. Jaffe, AS,, Ladapo, JA,; Lespe´rance, F, Medina, V, Newman, JD, Osorio, GA Parsons, F, Schwartz, JE, Shaffer, JA Shapiro, PA,. Sheps, DS, Vaccarino, V, Whang, W, Ye, S. Centralized, Stepped, Patient Preference–Based Treatment for Patients With Post–Acute Coronary Syndrome Depression CODIACS Vanguard Randomized Controlled TrialCODIACS Vanguard RCT. JAMA Internal Medicine, 2013, 1-8.

2. Ioannidis, J. P. (2005). Why most published research findings are false. PLoS Medicine.

This review required withdrawal for additional important reasons to the outdated search described by the editors, including inappropriate study inclusion, factually incorrect statements and conclusions not supported by the included evidence. As these reasons were not described in the reasons for withdrawal, we share a letter sent to the Cochrane Heart Group on March 30, prior to withdrawal, highlighting these issues:

**'Dear Cochrane Heart Group, In performing our own review and synthesis of the evidence for impact of diuretics on mortality and hospitalization in patients with heart failure, we identified critical issues in Cochrane review CD003838 that warrant immediate withdrawal of the review for revision.1 For the purposes of this letter, we focus on the analyses of 'mortality' and 'heart failure worsening' and 3 of the 4 randomized controlled trials (RCTs) that provided data for at least one of these analyses.2-4

First, the reviewers defined eligible participants as "adult participants with chronic heart failure, […] a clinical syndrome characterised by breathlessness and fatigue that is caused by an inability of the heart to support an adequate circulation, that may limit exercise tolerance and may lead to pulmonary congestion and peripheral oedema". Based on these criteria, reviewers should have excluded the trials by Burr, de Jonge, and Myers from this Cochrane review as none or few patients in these trials met this disease definition.1-3

In the trial by Burr et al,2 investigators excluded patients if "they had had congestive cardiac failure during the previous three months" or "they had ever had left ventricular failure". The published report further notes that "an attempt was made to discover the original reason for which each patient had been given a diuretic" and that "in most cases, however, this information was not in the hospital notes". They furthermore note that "ankle oedema was often mentioned in the notes, but in the majority there was no reference to cardiac failure". The trial by de Jonge et al specifically excluded patients with heart failure.3

First, the authors describe their objective as "to determine the effect of withdrawing diuretic drugs on oedema in patients prescribed them for only ankle oedema, excluding patients with cardiac […] failure". In their methods, the authors describe their approach to excluding patients with a clear diagnosis of heart failure, such as those "[having] congestive heart failure or increased risk of developing it after stopping diuretic drugs," or "heart failure previously established by a cardiologist, history of severe dyspnoea treated by the general practitioner as cardiac failure, atrial fibrillation, symptoms of right sided heart failure, palp-able right ventricular pulsations, or hepatomegaly". Therefore, de Jonge et al appeared to include only patients in whom heart failure was unlikely or ruled-out as the cause of ankle edema. Additionally, the Cochrane authors describe this trial as only including "participants with decreased ejection fraction (EF) measured by echocardiography", but this is not described in the published report of the de Jonge trial.1,3 Similarly, the eligibility criteria in the trial by Myers et al excluded patients with definite or probable heart failure.4

In this trial, "concurrent digoxin therapy" was the most common reason for exclusion, with additional relevant reasons for exclusion being "active heart failure […] (clinical or radiological evidence of heart failure)" or hyper-tension. Corroborating this is the fact that only 9 out of 77 included patients were noted to have had "previous CHF" according to the study report's patient characteristics table. Based on the above published trial details, it is clear that 3 of the 4 studies contributing data to the Cochrane review's 'mortality' and 'worsening heart failure' outcome analyses should be excluded according to the reviewers' predefined review eligibility criteria.1-4

Exclusion of the Burr and Myers trials leaves only the Sherman trial for the 'mortality' analysis, demonstrating no statistically significant difference between diuretics and control (0 versus 2 deaths). 1 Additionally, no trials remain for analysis for the 'heart failure worsening' outcome with exclusion of the Burr and de Jonge trials. As a result, exclusion of trials not meeting this Cochrane review's predefined eligibility criteria substantially changes the review's conclusions.

Second, further issues arise from attributing a causal role of diuretics in the reported reduction in mortality. In the trial by Burr et al, none of the 3 deaths in the control group were attributable to heart failure. 2 Similarly, only 1 of the 8 deaths reported in the control group of the trial by Myers et al was attributable to heart failure, with the others attributed to cancer (3), respiratory disease (2), stroke (1) or gastrointestinal bleed (1).4 Such inconsistencies between all-cause mortality and heart failure-related mortality would deserve, at minimum, description by the reviewers in their Discussion section. Ideally, the impact of inconsistency between outcomes should be considered in the determination of quality of the body of evidence, such as by using the framework provided by the GRADE approach as described in the Cochrane Handbook.5Third, the reviewers inappropriately exclude the results of the Myers trial from their analysis of 'worsening of heart failure'.1

According to the authors, they excluded Myers from this analysis "because of heterogeneity for heart failure worsening in the diuretic group versus placebo (chi-square, 16.03; P = 0.001)". The heterogeneity noted resulted from the increase in "withdrawal due to heart failure" in the diuretic group compared to the placebo group in this trial (6/29 versus 2/29). Such arbitrary exclusion from analysis constitutes selective outcome re-porting bias. Rather than excluding the trial by Myers et al, which contributed greater statistical weight (38.9%) than either trials by Burr or de Jonge, the appropriate course of action according to the Cochrane Handbook would have been to evaluate methodological and clinical sources of heterogeneity, and to abstain from performing a meta-analysis of this outcome. 5

Based on the above appraisal of critical issues present in Cochrane review CD003838, we urge editors of the Cochrane Heart Group to withdraw the aforementioned review from the Cochrane Library and issue a report on the Cochrane Heart Group website describing reasons for withdrawal. Authors of the review should then be provided with the opportunity to revise the review to meet the standards set by their protocol. Thank you for your consideration, Ricky Turgeon BScPharm, ACPR, PharmD; Michael Kolber BSc, MD, CCFP, MSc.

References:

1. Faris RF, Flather M, Purcell H, Poole-Wilson PA, Coats AJS. Diuretics for heart failure. Cochrane Database Syst Rev 2012;2:CD003838.
2. Burr ML, King S, Davies HE, Pathy MS. The effects of discontinuing long-term diuretic therapy in the elderly. Age Ageing 1977;6:38-45.

3. de Jonge JW, Knottnerus JA, van Zutphen WM, de Bruijne GA, Struijker Boudier HA. Short term effect of withdrawal of diuretic drugs prescribed for ankle oedema. BMJ 1994;308:511-3.
4. Myers MG, Weingert ME, Fisher RH, Gryfe CI, Shulman HS. Unnec-essary diuretic therapy in the elderly. Age Ageing 1982;11:213-21.5. Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

One commenter who responded to a post highlighted the controversy inherent in the type of open commenting PubMed Commons offered, “With all respect to the commenter, the misleading posted arguments and evident lack of insight underscores the danger of such unsolicited and unreviewed posting not subject to peer review.”

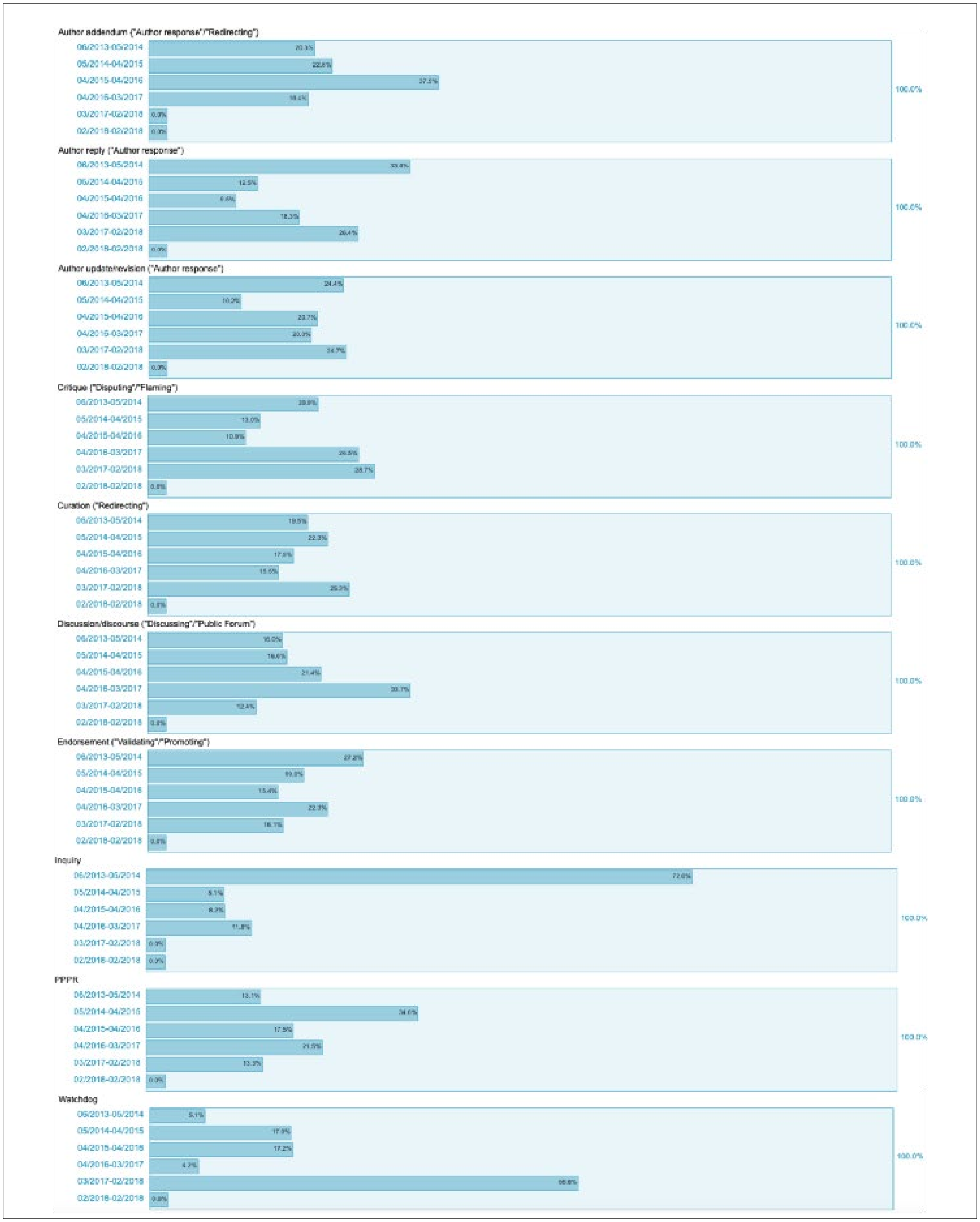
4.7.3 Associating Comment Content with Commenter Descriptors

Associating comment content with commenter characteristics and other descriptors in the Dedoose database offered insights into commenter personas. This type of analysis is much more granular than results reported from the the quantitative content analysis, especially when reviewing comment content that peaked during specific time periods. Figure 4.25 shows increasing popularity in certain comment content, as well as relative stability in comment content over the years.

A striking contrast was evident between commenters and the number of the PubMed-indexed articles they authored. The most highly published authors wrote 93% of response comments. Authors with the fewest number of indexed articles wrote nearly all the author update/revision posts. Commenters who had published a higher number of PubMed-indexed articles wrote posts that reflected academic writing considered publication-quality alternatives to scholarly communication (79%).

Figure 4.25

Changes in Comment Content While PubMed Commons Available



The more articles a commenter had indexed in PubMed, the more likely they were to write a critical comment or to ask a question. Alternatively, the fewer articles a commenter had indexed in PubMed, the more likely they were to write an endorsement of an article or to add further information about their own article (i.e., author addendum). Number of articles indexed in PubMed did not seem to affect the likelihood of a commenter engaging in a discussion. One hundred percent of watchdog comments were made by commenters who had written the least number of PubMed-indexed articles. Commenters who were the first author on a PubMed-indexed article constituted the majority of those who wrote a comment characterized as PPPR or as a reply to a previous comment (74%).

Regarding gender, 65% of watchdog comments were made by males; also, men posted more inquiry comments than women. More females than males redirected readers to other information resources. There was no variation between men and women who posted a comment characterized as an endorsement, PPPR, or a reply to a previous comment.

Commenting frequency affected comment content in unexpected ways. Commenters who posted more frequently than others on PubMed Commons were more likely to post an author reply, author addendum, or author update/revision. Eighty percent of all author replies were written by commenters who made more than six comments. Of the comments classified as alternatives to scholarly communication, 81% were made by commenters who made three or more comments. The more comments a commenter made, the more likely the post was to be a critique (e.g., 63% of critical comments were made by commenters who made 12 or more comments). Of the commenters who reported a correction or retraction, 70% had made between three and five comments. Commenters who made between six and eight comments were the most likely to curate information resources (36%); this group was followed by those who made between three and five comments (27%). Commenters who made one or two comments and those who made between 12 and 14 comments redirected readers at nearly the same rate (19% and 18%, respectively. Posting frequency did not affect endorsement comments.

Time of day a comment was made had little effect on comment content, although authors most often replied to a comment at 5 p.m. (14%), 8 a.m. (11%) or 3 a.m., 8 or 9 p.m. (10% each). Authors who updated or revised information contained in their article (i.e., addendum) often posted at 1 or 8 p.m. (19%) or at 8 a.m. (13%). Endorsement posts were usually written at 9 a.m. (11%), and inquiries were written at 11 p.m.

People in mid-career composed the majority of commenters who wrote critiques (49%) or whose posts were considered alternatives to scholarly communication (51%) or PPPR (52%). This was also the group most likely to provide an author reply (59%) or author addendum (37%). Early-career commenters were likely to curate content (42%), write an endorsement (59%), or ask a question (44%). The most senior authors often updated or revised information about their articles (52%) and infrequently curated content (27%) or wrote a watchdog comment (16%).

Commenters who made comments on PubMed-indexed articles that were published in journals that also published their own work could, arguably, be considered an authentic peer. For these individuals, comment content was typically an author addendum (90%), reply (94%), or update/revision (71%). These commenters were less likely to write a critique or inquiry, endorse an article, or redirect readers. An overwhelming number of watchdog posts (83%) were made by this group of peers.

4.8 RQ1d: Types of Communities Forming on PubMed Commons

Social Network Analysis. The increasing awareness of social networks and their influence on individual behavior, organizational culture, and social movements (Christakis & Fowler, 2009; Duhigg, 2012) sparked my interest in performing a social network analysis (SNA) of PubMed Commons. It's widely accepted that, "the most important feature of Internet forums is their social aspect" (Morzy, 2013, p. 623). Iterative data collection and analysis of PubMed Commons since its launch in 2013 suggested that the forum network was fragmented with few recurring group interactions. My quantitative analysis of commenters verified that only a handful of commenters regularly commented on the same articles or consistently interacted with one another on an ongoing basis. Performing a holistic SNA of the forum provided evidence that few groups of people who shared a concern or passion about a similar topic (i.e., communities of practice [Wenger, 2000]) were engaged in regular scientific discourse on PubMed Commons. An SNA of PubMed Commons provided a necessary third lens for my explanatory case study research.

SNA has its origins in social science and network analysis/graph theory. Much of SNA development comes from mathematicians, physicists, biologists, and computer scientists, each studying networks of different types. Widespread advancements in computing have made it easier to apply SNA to a range of problems. Social scientists use SNA to better understand how a network functions and how to improve its effectiveness. One advantage of SNA is visualization of the data to uncover patterns in relationships. Using SNA to study computer mediated communication (CMC) helps identify causes for dysfunctional networks and to promote social cohesion and growth in an online community. Social network sites like Facebook use basic elements of SNA to identify and recommend potential friends based on friends-of-friends. Network operators use SNA methods to optimize the structure and capacity of their networks.

SNA is a methodology for examining networks. This research approach provides a vocabulary and set of measures for studying relationships. Several SNA theories posit how a network functions based on statistical measures and visualizations of the data. A social researcher's perspective focuses on the structure and shape of a collection of relationships (i.e., ties, connections, interactions, links, edges) between individuals (i.e., nodes, entities, alters, vertices). For example, an investigation could examine what and how a factor under investigation (e.g., friendships, money, ideas, power, disease) flows through the network. The unit of measure is the connections embedded in the network, not the characteristics of individuals. When collections of relationships are analyzed, network patterns become evident. The variety of shapes and patterns (e.g., fragmented, divided, unified, spoke, cluster, clique) aids in identifying key individuals located in important positions of the network (e.g., hub, bridge, gatekeeper, island).

Unlike real-world networks that occur randomly, social networks are formed based on individuals' choices and actions. For example, the PubMed Commons social network (PMCSN) was formed when individual commenters made decisions about which article(s) warranted their time for making a comment and whose comments they should reply to.

A one-mode, nodelist, 913-cell x 913-cell Excel spreadsheet was developed from the list of individual network commenters who was one of at least two commenters on any given PubMed-indexed article. Once again, data was pulled from MS Word documents created from article metadata and corresponding comments that were copied/pasted from the PubMed GUI (following a PubMed search for “all[sb]”) before comments were deleted from public view. Ordinal numbers (0, 1, 2, 3, etc.) corresponded with the number of times a specific commenter commented on the same article with another specific commenter. These tallies were entered into the matrix cells so that UCINET algorithms could calculate the connectedness of the overall forum network and the relationships between commenters.

A NetDraw program processed statistical calculations into visual representations of commenter relationships (i.e., sociogram), which depicted a) who interacted with whom (i.e., posting on the same article), b) the strength of relationship between commenters (i.e., how often commenters posted on the same article), and c) directional flow of communication (i.e., who commented on whose comments). Results from both types of data analysis are detailed below in two subsections, namely statistical measures and sociogram visualization.

4.8.1 Statistical Measures

The Excel spreadsheet described above was saved as a DOS file and imported into UCINET for Windows (Visit https://sites.google.com/site/ucinetsoftware/home for more information about the program). The DOS file was then saved as a one-mode, nodelist adjacency matrix in dl format. Various multidimensional scaling analyses of matrix data were performed using standard UCINET queries, and dynamic filters were applied to remove ties with low closeness centrality (i.e., those not closely connected), which were peripheral to the cohesiveness of the network (i.e., weakly connected to the central network). Figure 4.26 shows cohesion measurements for the PMCSN. Connectedness was only 0.054 on a normalized scale of 0 to 1. “The idea of cohesion is connectedness or knittedness” (Borgatti, Marten, & Johnson, 2013, p. 150). Readers should note the “Network is disconnected” warning at the bottom of the figure.

Figure 4.26

Network Cohesion Statistics for PubMed Commons Social Network

```
NETWORK COHESION
--------------------------------------------------------------------------------

Input dataset:                          UCInet PMC (C:\Documents and Settings\Michelle\Desktop\UCInet PMC
Output dataset:                         UCInet PMC-coh (C:\Documents and Settings\Michelle\Desktop\UCInet PMC-coh
Ignore direction of ties:               NO (C:\Documents and Settings\Michelle\Desktop\NO
Ignore reflexive ties:                  YES (C:\Documents and Settings\Michelle\Desktop\YES

Measures

                                1
                          -------
    1         Avg Degree    1.742
    2            H-Index       11
    3            Density    0.002
    4         Components      512
    5    Component Ratio    0.560
    6      Connectedness    0.054
    7      Fragmentation    0.946
    8            Closure    0.371
    9       Avg Distance    5.823
   10        SD Distance    3.064
   11           Diameter       20
   12            Breadth    0.987
   13        Compactness    0.013

13 rows, 1 columns, 1 levels.
WARNING: Network is disconnected. Distance-based measures are calculated within components.

-----------------------------------------
Running time: 00:00:01 seconds.
Output generated: 18 Dec 08 06:47:09
```

In social network analysis, the term network cohesion refers to a measure of the connectedness and togetherness among actors within a network. A social network can be defined as a network formed by a set of interacting social entities (actors) and the linkages (relations) among them. The index of network cohesion is a single value that captures the togetherness of the group. Network cohesion can be measured in a variety of ways, most of which are based on the dyadic cohesion (or closeness) between a pair of actors. This measure should be differentiated from closeness centrality, which measures how close an actor is to all other actors within a network.

The simplest measure of network cohesion is density (i.e., the number of relationships in the network). A perfectly connected network has a density of 1. The PMCSN density measured extremely low at 0.002 with standard deviation of 0.051, which signaled a low number of connections between commenters. In low density networks, information does not transmit very efficiently because it must flow from member to member, rather than diffusing from one member rapidly to all the others. Another issue with low density networks is the "hit by a bus" problem, where if one or two members are taken out of the network, the remaining network breaks down because those members are no longer there to coordinate the different members who don't talk to each other. Denser networks are less vulnerable to disruption when key members leave.

The density problem becomes exacerbated with network size. The PMCSN diameter (i.e., number of steps necessary for information to flow from one side to the other) was 20, indicating the network was indeed large.

Another social network measure—average degree—is not dependent on network size, which makes interpreting its significance easier. "It [average degree] is literally the average number of ties that each node has" (Borgatti, Marten, & Johnson, 2013, p. 152). For example, average degree in the PMCSN was 1.742, which means that on average each commenter posted with some other commenter on the same article about 1-3/4 times. This number further strengthens findings from the quantitative content analysis and Dedoose results reported above, which both reported that two people usually commented on any given article.

Like density and average degree measures, component and component ratio measures are relative, with component ratio easier to interpret because it is a normalized measure with a maximum value of 1—where every node is an isolate—and minimum value of 0—where the network is one large component (Borgatti, Marten, & Johnson, 2013, p. 153). Larger values indicate less cohesion, although the result is not sensitive and cannot be evaluated without comparing it to another similar network or the same network at another snapshot of time. Component ratio for PMCSN was 0.560.

More revealing measures are connectedness (i.e., proportion of node pairs that can reach each other by a path of any length) and fragmentation (i.e., number of node pairs that cannot reach each other by any means). Obviously, these two measures are inversely related, as evidenced by the results of UCINET calculations: connectedness was 0.054, and fragmentation was 0.946. Like component ratio, when taken out of context or without comparing results to a similar network, the measures are meaningless. However, changes in these measures can be used to evaluate a network in a what-if simulation (Borgatti, Marten, & Johnson, 2013, p. 154).

Compactness is a variation to connectedness and fragmentation. This measure has a value of 1 when the network is a clique (i.e., all commenters directly interact with one another) and 0 when the network is entirely made up of isolates (i.e., commenters have no interactions). A compactness measure of 0.013 is very low and indicates that PubMed Commons was a *community of independents* who generally commented in isolation.

Although measure of compactness is more indicative of network cohesion than distance, the latter is an important macro-characteristic of a network (Hanneman, n.d.). Distance is the number of relations configured in the shortest possible steps from one node to another. A PMCSN distance of 5.823 implies that making connections was troublesome and slow. Readers can relate this finding to Kevin Bacon's popularization of sociologist Stanley Milgram's (1967) "six-degrees of separation." PMCSN commenters had six-degrees of separation.

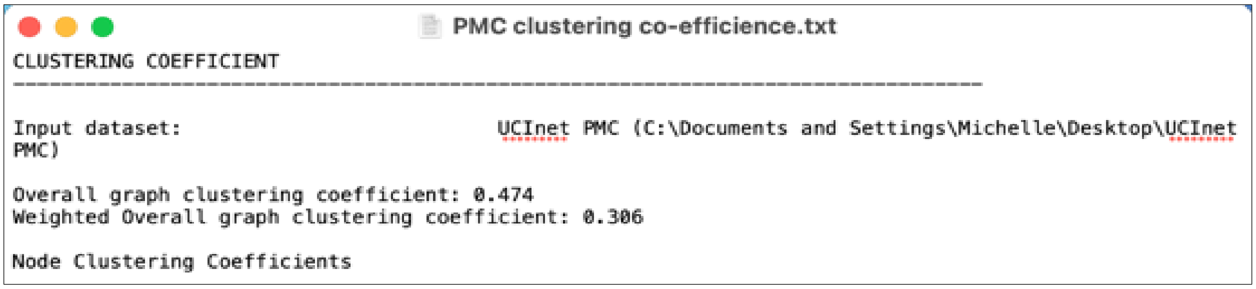
Network closure (i.e., transitivity) can help better explain these difficulties, as it conveys completeness of relational triads. For example, in social relations, we expect that if A knows B and B knows C, then there would be a relationship between A and C (Borgatti, Marten, & Johnson, 2013, p. 155). In the PMCSN this would imply that if commenter A posts with commenter B on an article, and commenter A also posts with commenter C on an article, there is a tendency for commenter B to post with commenter C on an article.

Small group theorists posit that "all of the really fundamental forms of social relationships can be observed in triads" (Hanneman & Riddle, 2005, para. 23). Because of this, it is suggested to conduct a triad census for the network as a whole. In PMCSN, total number of all possibilities that this type of relationship would happen was 758,549,616. Transitivity measures the number of closed triplets in a node's neighborhood over the total number of triplets in the neighborhood. The transitivity percent of triples in which A>B and B>C that are transitive in PMCSN was 37.09%. The transitivity percent of triangles with at least two ties that were closed with a third tie was 14.27%. Without comparison to another similar network, this measure is again not terribly informative, except for noting the potential for connections was greater than achieved.

The weighted overall graph clustering coefficient of PMCSN was .306 (see Figure 4.27). Watts and Stogatz (1998) found that networks with low step distance and high clustering coefficient are highly connected. Conversely, PMCSN had a high step distance and low clustering coefficient, which is yet another indicator that the network was not connected.

Figure 4.27

Low Weighted Clustering Coefficient Indicates Network Potential Not Met



```
PMC clustering co-efficience.txt
CLUSTERING COEFFICIENT
--------------------------------------------------------------------------------

Input dataset:                     UCInet PMC (C:\Documents and Settings\Michelle\Desktop\UCInet
PMC)

Overall graph clustering coefficient: 0.474
Weighted Overall graph clustering coefficient: 0.306

Node Clustering Coefficients
```

When comparing the overall network clustering coefficient (i.e., number of closed triplets of nodes) with individual measures for commenters who either posted most often or posted with a high number of other commenters on the same PubMed-indexed article (see Table 4.14), the difference between what could have been a successfully cohesive network and the PMCSN becomes obvious.

Table 4.14

Most Connected Commenters with Low Clustering Coefficient and High Pairs

Commenter	Clustering Coefficient (0< and >1)	Pairs (>100)
Bastian, Hilda	0.024	861
Bates, Clive	0.036	253
Bishop, Dorothy	0.068	190
Bramer	0.103	136
Corcos, Daniel	0.392	153
Keller, David	0.020	153
Maniatis, Lydia	0.068	406
Puliyel	0.343	105
Southan, Christopher	0.038	120

The same type of comparison can be made when looking at betweenness measures for the network and top individual commenters. Betweenness measures show how likely a person is to be the most direct route between two people in the network (i.e., who is the person through whom most of the information is likely to flow and who has significant influence—good or bad—over what flows). These individuals serve as “brokers” in the network.

Figure 4.28

Betweenness Data for Most Connected Commenters

PMC betweeness.txt

```
Important note: This routine cannot handle valued data, so it binarizes your data automatically.
                It DOES handle directed (non-symmetric) data, so it does NOT symmetrize.


Un-normalized centralization: 24701324.427

                                             1            2
                                  Betweenness nBetweenness
                                  ------------ ------------
  45                    Bastian     27292.834        3.285
 156                  Colquhoun     16260.833        1.957
  61                     Bishop     14948.833        1.799
 624                    Oransky     12002.000        1.445
 250                     Fenner      9487.500        1.142
  86                     Brembs      9377.000        1.129
  84                     Bramer      9336.000        1.124
  46                      Bates      8787.000        1.058
 887                   Woodgett      8004.733        0.963
 763                    Southan      6589.600        0.793
 351                      Hegde      5296.700        0.638
  52                     Berger      4838.000        0.582
 820                   Turner E      4378.000        0.527
  82                    Boutron      4251.000        0.512
 418               JC Geriatric      4231.000        0.509
 553                   Masukume      3256.600        0.392
 879                Willighagen      3039.667        0.366
 139                   Chiolero      2928.000        0.352
 806                 Tibshirani      2850.500        0.343
 416                   JC CREBP      2597.000        0.313
 422                    JC Neph      2366.000        0.285
 659                    Puliyel      2299.200        0.277
```

Figure 4.28 shows a familiar list of active PMCSN commenter names. Betweenness mean for PMCSN was 237.714 and normalized (n)Betweenness was 0.029. Note that Bastian’s betweenness (27,292) and nBetweenness (3.285) were over 100 times the mean for the network. The most average commenter on PMCSN was Melissa Greenwald who made two comments on two different articles and interacted with three other commenters: 1) Preen Berthelsen (who authored 21 posts on 18 articles—one on which Greenwald also commented); Claus U Niemann (who made one post on one article that Greenwald also commented on); and The Neph Journal Club (who made 30 posts on 30 articles—one of which Greenwald and six others commented on).

Tables and figures highlighting PMCSN leaders (i.e., network champions) show that regardless of method or tool for data analysis, the same people are identified as key individuals in the network.

4.8.2 Sociogram Visualization

A global network sociogram (i.e., visual map showing overall network patterns and shapes) was created by importing statistical information from UCINET into the NetDraw visualizing program. The layout of a network diagram refers to node position in the diagram. Notably, the value of graph layouts, like those generated using NetDraw, is the pattern showing which nodes are connected to which others by placing them in such a way that the connections are easier to see (Borgatti, Marten, & Johnson, 2013, p. 101, 105). Macro- and micro-level analyses were performed on the PMCSN. Filters were used to show clusters of commenters. Macro level analysis demonstrated the criticalness of the social ecology of the limited number of smaller communities, especially three-legged closed triangles, that formed. The micro-level (i.e., sociocentric) view showed the social composition of interactions between commenters, types of relationships, paths between commenters, and connectedness (i.e., density) of commenters. Figure 4.29 shows the unfiltered network and highlights the intensity of relationships (i.e., ties indicated by lines) between commenters (i.e., nodes indicated by blue squares).

Figure 4.29

Unfiltered Network Sociogram

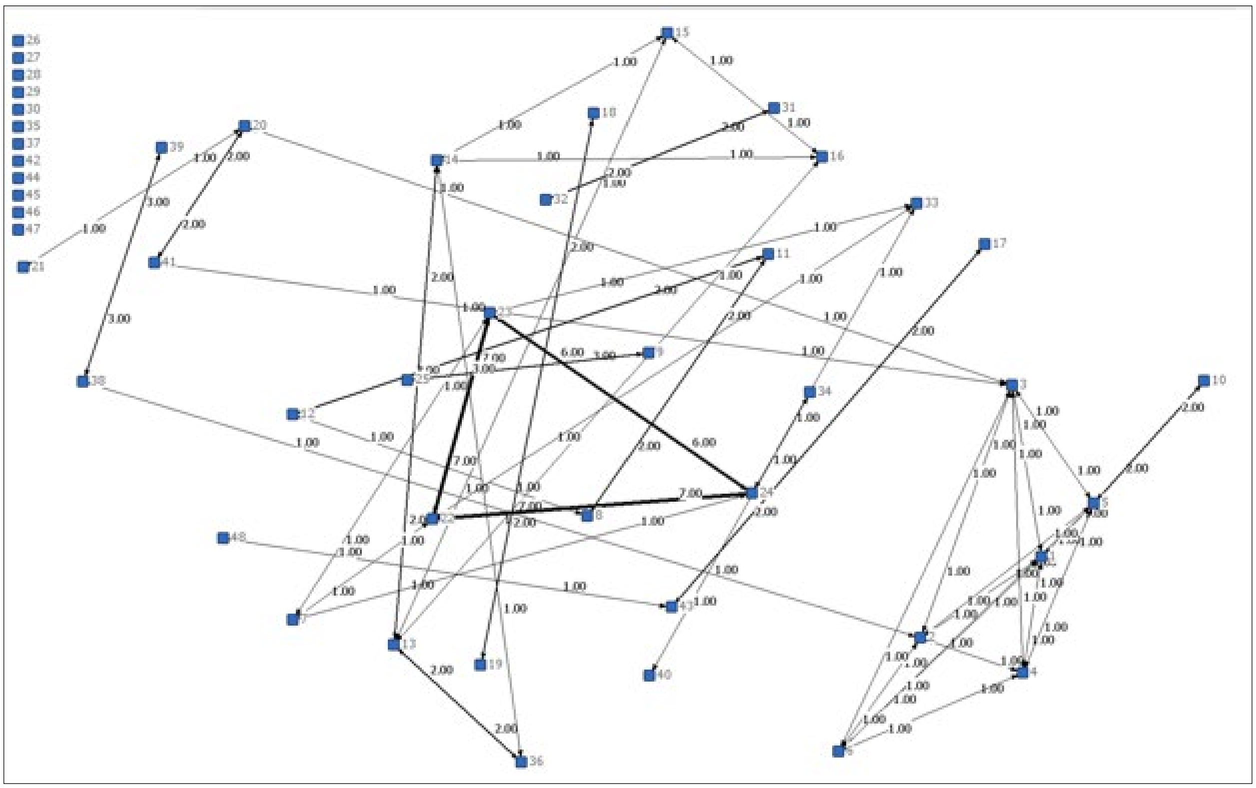
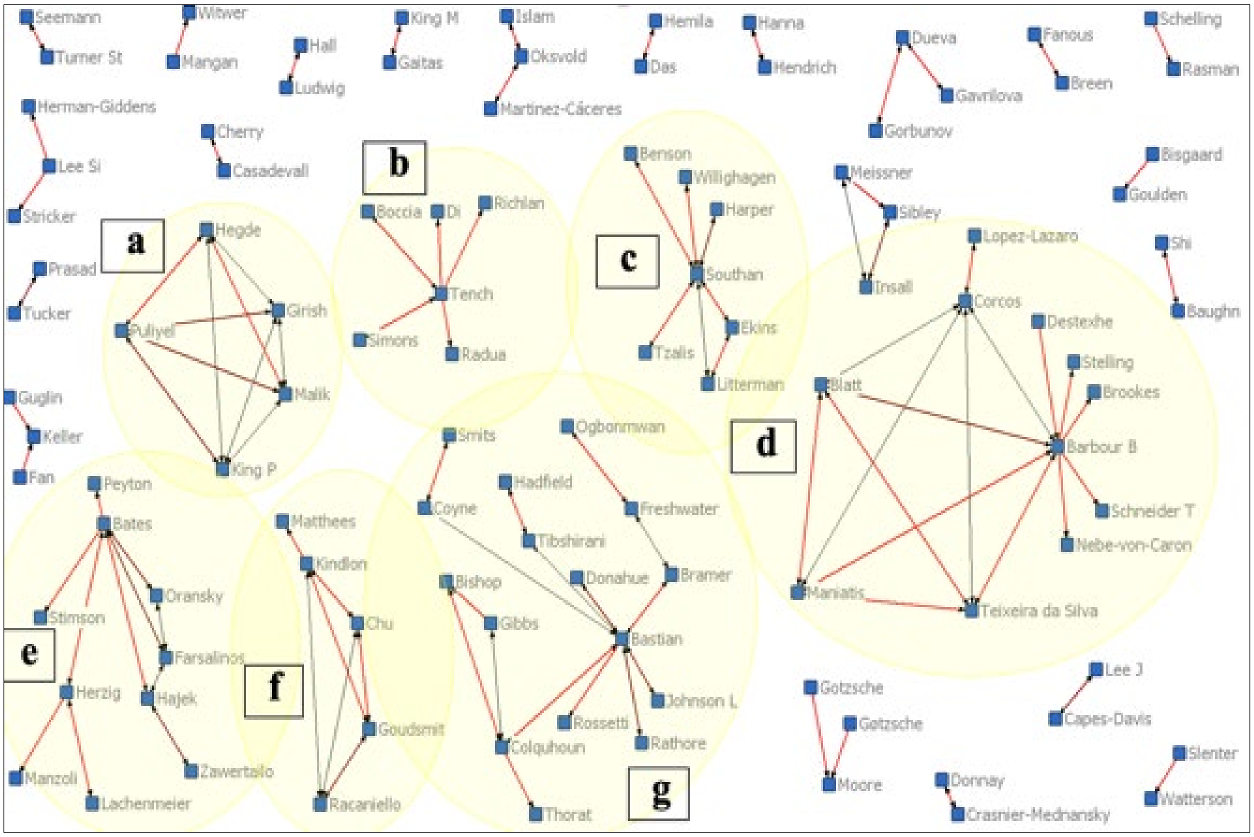


Figure 4.30 clearly demonstrates that the PMCSN was a fragmented and disconnected network with few relationships among commenters. The figure shows 15 edge (i.e., two nodes connected with a tie) and six 2-star groups (i.e., three nodes connected with two ties that left a “structural hole”) that are positioned around seven larger clusters composed of five or more nodes and ties between them. A cluster is a collection of nodes (i.e., individuals in PMCSN) with dense relationships internally and sparse relationships externally. The nodes are positioned closer to each other than to other nodes.

Figure 4.30

Filtered Network Sociogram to Emphasize Commenter Clusters



In Figure 4.30, clusters are highlighted in yellow and labeled with lowercase letters “a” through “g.” It is important to note that the sole unlabeled transitive triad (i.e., closed triangle or “ring” network) is positioned between highlighted clusters “c” and “d”. The composition of Figure 4.30 is the result of filtering NetDraw visualized data to focus attention on clusters and to display relationships that were indicative of those throughout the entire network.

Line color between connecting commenter-labeled nodes on Figure 4.30 adds further information about relationships formed on the PMCSN. Black lines indicate when connected commenters posted once on the same PubMed-indexed article. Red lines indicate when the connected commenters posted two or more times on the same PubMed-indexed articles.

Thicker lines in both Figure 4.29 and Figure 4.30 indicate more frequency in comments made on the same article. Directional arrows in both figures indicate flow of comments. For example, the top right connection in Figure 4.30 shows that Schelling commented on Rasman's post, yet Rasman did not comment on Schelling's post. A two-headed arrow connecting nodes indicates that comment posting was reciprocal. Commenters are identified and cluster shapes/structures are named in Table 4.15.

Table 4.15

Significant Individuals and Structures within PubMed Commons Network

Network	Nodes	Significant Individual(s)	Network Community Shapes and Structures
b	6	Trench: Hub	Mixed Hub-and-Spoke
c	7	Southan: Hub	Cluster with Bridge
d	11	Corcos: Gatekeeper Barbour: Hub	Kite with High Density Clique and Bridge to Cluster
e	10	Bates: Gatekeeper	Kite with Bridge connecting Cluster and Clique
f	5	Kindlon: Gatekeeper	Kite with High Density Clique
g	16	Bastian: Hub Colquhoun: Gatekeeper	Cluster Bridged with Clique

PubMed Commons formed an overall unified, rather than divided, network with individuals forming a few small- to medium-sized clusters. A divided network is created when individuals form two apparent groups. The sociogram in Figure 4.30 shows that cluster "a" is a complete cluster comprised of five nodes that each have a tie to every other node. It is also a highly dense clique, with all relationships present among themselves. Cluster "b" is a star or spoke graph with Trench centrally located as the ego or hub of the network and facilitating connections between all other commenters in the cluster. Similarly, Southan is the hub of the cluster "c" star graph, although connections with both Ekins and Litterman complete a closed ring group. Like cluster “a,” cluster “d” forms a complete clique graph with five commenters. Barbour acts as a hub in the center of the star cluster, acting as a bridge between the star and the clique. Corcos also serves as a bridge and a gatekeeper to Lopez-Lazaro. Corcos alone connects Lopez-Lazaro to everyone else in the cluster. Of interest is that one path option between Lopez-Lazaro and Brookes exemplifies the average distance of the PMCSN, which equals six steps (>): Lopez-Lazaro > Corcos > Blatt > Maniatis > Teixeira da Silva > Barbour > Brookes.

In cluster “d”, Bates is the bridge and gatekeeper, as well as the hub between several commenters, some who formed ring groups and others who also took on a bridge role. Like Bates, Kindlon is the bridge and gatekeeper for cluster “e.” Without Kindlon, the others in the “f” clique would be isolated and not have a relationship with Matthees. Cluster “g” centers around hub Bastian who, according to all data analyses for the PMCSN, served as the linchpin to many relationships. She is the path between many individuals who might not otherwise be connected.

Individuals who serve as hubs, brokers, and gatekeepers are typically the most central or connected individuals and have the highest number of relationships with others. Often, they link people that have no other individuals in common, and their position between individuals facilitates the shortest step to all others. Preferential attraction to individuals is affected by popularity (e.g., people want to be associated with popular people, ideas, and items, thus further increasing their popularity) and quality (e.g, higher ‘quality’ individuals will naturally attract more attention, faster). Individuals who reach critical mass first become champions for many individuals and followers. Higher embeddedness (i.e., number of common neighbors) leads to more trust in an individual.

The purpose of a social network analysis in online social media forums, like the one detailed in this subsection, is to leverage automated systems to “identify more or less collaborative contributors, assess [network] community health, and decide where interventions or support might be most helpful” (Hansen, Shneiderman, & Smith, 2011, p. 264). A similarly disconnected PMCSN was also identified by Farabough and Burns in 2015 (see Appendix D). Chapter 5 of this dissertation will further detail utilizing SNA for evaluating online forums and making recommendations for changes based on findings.

4.9 RQ 2: Factors Affecting Scientific Discourse on PubMed Commons

I synthesized the data collection and analyses detailed above with my ongoing systematized search for information from the literature and social media about PubMed Commons to answer RQ2: What factors contributed to or limited the scientific discourse on PubMed Commons? Themes that were affecting the adoption and use of the forum emerged during the process of constantly comparing results from my primary data with perceptions about the forum that were articulated by PubMed Commons’ organizers, early adopters, and potential adopters in documented publications, gray literature, and social media content (e.g., blogs and their comments, tweets, YouTube videos).

I used the same grounded theory procedures I used for the qualitative content analysis of PubMed Commons’ comments to arrive at factor categories. In this way, factors contributing to or limiting discourse on the forum were identified, named, categorized, and described. Emerging themes were considered during open coding, and categories were repeatedly reviewed to the point of saturation (i.e., no new factors were detected).

Table 4.16

Focused Category Factors Contributing To or Limiting Scientific Discourse

Focused Category	Characteristics
Anonymity	• Commenters had to identify themselves. • Commenter names were visible on the PubMed GUI. • Names remain associated with comments on the commons_archive.csv file. • Individuals who shared opinions about PubMed Commons were not PubMed Commons commenters.
Alternative commenting platforms	• PubPeer, F1000, Retraction Watch, arXiv, biorXiv, Cochrane Crowd/Task Exchange, Research Gate. • Individual journal forums/clubs. • Blogs and Twitter accounts.
Rules of engagement	• Commenter must be author of an article indexed on PubMed. • Problems with sign-up; difficulties registering. • No clear or explicit definition of writing expectations. • "Discourse" and "discussion" were not defined; both terms were used interchangeably by PubMed Commons organizers. • Initially, no explicit mention of PPPR from PubMed Commons organizers; term began creeping into the narrative as time went by. • No stated purpose for the forum was communicated. • 800-word limit to posts (originally,) but commenters got around this by making successive comments.
Moderator involvement	• After going through an automated filter, comments were immediately uploaded to the PubMed GUI. Moderators subsequently reviewed visible comments and removed those that violated guidelines. • Forum was promoted as not moderated; however, comments were annotated on the PubMed GUI as "removed by moderator," "edited/deleted by author." • Moderating was not transparent; commenters who made comments violating guidelines were offered to edit the comment or have the comment removed. • Several commenters blogged about their experience of having comments pulled and being banned from the forum. • No visible time stamp on comments.

(table continues)

Focused Category	Characteristics
Global scope of forum	• PubMed is the world's largest database for indexed biomedical and life sciences articles; over 28 million at the time PubMed Commons was functional. • Daily traffic was 2.5 million global users; 3 million searches; 9 million page views. • There is no other single, centralized source with both a built-in audience and a platform for the forum. • Accessing PubMed was already built into the daily workflow of researchers, librarians, and practitioners. • Access to PubMed GUI is free. • Anyone with connectivity to the Internet can see PubMed Commons comments.
Little promotion	• Organizers and editors relied on NCBI Insights and PubMed Commons Blog to communicate news and forum updates. • NCBI and NLM occasionally hosted webinars and made various presentations. • Organizers published only one peer-reviewed article focused on PubMed Commons. • PubMed Commons logo, trending articles, and search filter for "Reader Comments" was placed onto the PubMed GUI well after the forum was open for commenting. • PubMed Commons had a Twitter social media account but did not utilize other social media for promotions.
Outdated technology	• Findability and searchability of comments was difficult before the "Reader Comments" filter was available. • API was not functional for interfacing with other systems. • Technological functionality lagged other similar alternatives. • No Google extensions. • No doi for comments. • No user ID to disambiguate username. • Archived file did not preserve helpful/not helpful comments or threaded/nested commenting.
Low participation; few comments	• Journal editorials, conference presentations, blog posts, and social media content by individuals interested in PubMed Commons shared ongoing concerns about low participation. • PubMed Commons editors indicated participation was not important; instead, having an open space available for commenting on publicly funded research was purportedly most important. • Discontinuation was attributed to low participation and other commenting alternatives.
Interactive communities not forming	• Most posts were single, not threaded. • Articles with the most comments had few participants. • Only a few commenters posted regularly.

(table continues)

Focused Category	Characteristics
	• The network was disconnected. • Differences in online and offline connections affected motivation to contribute.
Strong emotions about PPPR	• Most literature and social media focused on PubMed Commons as a space for PPPR, not simply discussing articles that had been published. This is a big distinction. • Two articles with most comments attracted commenters who were passionate about PPPR and commenter anonymity. • Social media comments questioned what constituted a "peer," asking whether every author of any article indexed on PubMed qualified the to be a peer in every field. • PubMed Commons editors focused an increasing amount of attention on PPPR issues (e.g., erratum, expressions of concern) and less attention on promoting the forum.
Lurkers, non-adopters, and members who did not post	• It is difficult to measure the value of the forum for individuals not making a visible contribution on the forum. • Organizers did not report data about how unidentified participants/members were using information they read on PubMed Commons. • Lurkers are important members of an online community, as they might be sharing forum information in their own social circles and on social media.
Sponsorship	• Who was funding the project? • Other than the editors, who else was involved in the project? • What priority was given to technological needs? • Editors were having to manually collect and analyze data about the forum. • Who was managing the project?
Time return-on-investment (ROI) for commenters	• What was the benefit of commenting to commenters in terms of social capital, attention, and/or career advancement? • It is widely known that busy researchers, scientists, and practitioners have little time to do much of anything outside of regular work. • Organizers did not promote benefits of commenting to potential adopters or to forum members/participants. • Incentives line up with traditional reviewing (e.g., none other than helping shape the field; relieving guilt that others are reviewing your work; networking; helping out a friend; you might get a thank you for reviewing).

(table continues)

Focused Category	Characteristics
Premature discontinuation	• NCBI Insights announced the abrupt discontinuation of the forum with only a month's notice and with very little explanation. • Editors could not discuss the forum with me until after a post-mortem review with organizers. I was unable to make a connection with them for a follow-up interview even though initial emails were exchanged. • There was a lot of mass communication and social media buzz about the discontinuation; many shared the belief that the forum closed prematurely.

In the focused coding phase, properties and dimensions of certain categories were linked, and subcategories were created in others. Focused categories included: anonymity; alternative online spaces for after-publication commenting; forum guidelines; moderator involvement; global scope of potential adopters; lack of forum promotions; technology; small number of comments posted; lack of community forming; strong emotions about PPPR; questions about sponsorship; time return on investment for commenters to write and post comments on PubMed Commons; and premature discontinuation (see Table 4.16). These factors are briefly operationalized below . Final categories developed during the axial coding phase are fully explained as part of my discussion in Chapter 5, and each is supported by evidence from my primary data collection and from information found during my systematized search.

4.9.1 Grounded Practical Theory and Inductive Theory Building

The final stage of grounded theory is axial coding, which involves discovering interrelationships among codes. Recall in Chapter 3 that I referenced Creswell (2007) who advised that discovering interconnected themes is a necessary part of rigorous research. I also mentioned that I would employ grounded practical theory (Craig & Tracy, 2009) to characterize real-world lessons learned about PubMed Commons and the phenomenon of an online forum aimed at biomedical and life sciences researchers for commenting about published articles. Fulfilling these promises serves as the basis for Chapter 5. In the next chapter I explain how I used inductive reasoning to link my data analysis with the factors that affected the use/disuse of PubMed Commons. I developed the forum innovation agility model to assist other organizations who are interested in introducing a commenting system like PubMed Commons.

4.10 Chapter 4 Summary

Chapter 4 provided a comprehensive overview of my seven-year investigation of PubMed Commons. Mixed methods validated findings obtained through different lenses. The first section of this chapter detailed the process for data collection, the way in which datasets were organized, and various tools used for data analysis. The second section provided primary data results and analyses that were necessary to answer RQ 1—How did PubMed Commons function as an online forum for posting comments about published articles? Subsections answered four additional parts to RQ1. Methodologies for investigation included: 1) CMDA (i.e., quantitative and qualitative content analysis), 2) the creation of a database using Dedoose software, 3) a social network analysis, and 4) a systematized search for journal articles, gray literature, and social media content about PubMed Commons. A third section of this chapter answered RQ2 by providing a list of 14 factors that contributed to and limited discourse on PubMed Commons. These categories emerged following focused coding of a grounded practical theory analyses of the forum. Findings from this final phase of axial coding serves as the basis for discussion and inductive theory building in Chapter 5.

CHAPTER 5

DISCUSSION AND CONCLUSIONS

The iterative nature of the seven-year investigation (March 2014 to June 2021) reported in this dissertation pointed to confounding factors that could have affected the adoption of PubMed Commons and thwarted its ability to successfully function as a platform for "open discourse about published articles" (Collins & Tabak, 2014, p. 613). Chapter 5 is scaffolded with a discussion, recommendations, limitations, suggestions for future research, and concluding thoughts. A systemic overview of the data collection and analyses reported in Chapter 4 was tempered with peer-reviewed articles, grey literature, and social media posts about PubMed Commons. An ongoing environmental scan of similar forums grappling with post-publication commenting provided necessary background information for an intensifying social movement that addresses peer review in scholarly communication. A key informant interview with PubMed Commons co-editors Bastian and Vaught added beneficial personal insights from two individuals who were hired as contractors by NCBI to steward the PubMed Commons forum. Blending these data sources, this single case study posits 11 factors that could have affected the adoption of PubMed Commons and might have influenced its discontinuation.

To arrive at insightful generalizations, a reflective, explanatory case study must also consider current events occurring at the time of investigation. Regarding PubMed Commons, these external influences included emergent networked information communication technologies (ICTs), escalating adoption of computer-mediated communication (CMC), and evolving publishing models and metrics (i.e., altmetrics).

The dramatic increase of influence by online forums is changing many aspects of our lives, from how we decide what to buy or where to travel to how we manage our health. Its effects on scholarly communication are poised to democratize institutions entrenched in traditional peer review and upend society in consequential ways like those following the 15th century printing revolution (Aluetta, 2010; Naughton, 2012). Like the moveable-type press, the internet is outfitted to facilitate communication and knowledge sharing more broadly and rapidly than humans have ever experienced. Considering such complexities, the discussion section of this chapter explains how results from the quantitative/qualitative content analyses and social network analysis detailed in Chapter 4 overwhelmingly point to applying Rogers' Diffusion of Innovation (DOI) Theory to frame lessons learned. In this chapter, I am introducing a model grounded on Rogers' original DOI model and informed by my case study, which incorporates today's more complex socio-technological communication environment. In my forum innovation agility model, 11 factors affecting PubMed Commons adoption are organized in four layers that flow from inside to outside locus of control. Recommendations for organizations implementing a similar online forum conclude the discussion section of this chapter.

5.1 Diffusion of Innovation Theory

Heralded as "a pilot commenting system for authors in PubMed" (NCBI staff, 2013a), PubMed Commons' globally accessible platform was the first-of-its-kind to offer a singular, large-scale, open space for authors to comment on an ever-growing collection of biomedical and life sciences literature. During the short lifecycle of the forum, some 21.5 million database records grew to 27.8 million (NLM, 2015; NLM, 2019). Influenced by the rising popularity of interactive Web technologies, scientists voiced a desire and need for an open online venue to post comments about published works. Only a few small enterprises had attempted this feat in advance of the PubMed Commons initial, invitation-only, soft pilot launch in June 2013 (Dayton, 2006; Katz & Redberg, 2018; Lizarondo, Kumar, & Grimmer-Somers, 2010; Regenburg, 2010; Tibshirani, 2013). As early as 2006, website forums like *BioWizard*, which hosted comments only on articles reached by searching PubMed, and *JournalReview.org* attempted to aggregate post-publication commenting among disciplines. Publishers and editors (e.g., *PLoS ONE* and *BioMed Central*, including *Retrovirology, PLoS Biology,* and *Cell*) tested the post-publication commenting waters by launching journal-specific online reader commenting tools (Dayton, 2006).

In 2011, an interested group of scientists approached leaders at NCBI with an idea to host a commenting and rating forum like those available on Amazon and IMDb (Internet Movie Database). Together, the group would embark on a journey to guide ICT and CMC innovation. With the creation of a social mass communication channel for researchers, the group was introducing a novel method for promoting a culture of what Collins and Tabak (2014,) described as a process for self-corrective research that might rescue the “hobbled, short-term checks and balances system blamed for compromising research reproducibility”(p. 612). With all good intentions, something happened on the way to the forum.

Peer networking facilitated by the forum is important for diffusing innovative technologies into an established institution (Kaminski, 2011). For example, innovators, like the scientists who spearheaded PubMed Commons, and early adopters, like commenters who initially commented on the forum, would expectantly serve as change agents (i.e., project champions) to “influence their peers through peer-to-peer communication, role modeling, and networking" (p. 4). French sociologist Gabriel Tarde hypothesized a process for leveraging peer imitation for social change as early as 1903 (Kinnunen, 1996). Rogers further developed concepts about innovation diffusion in 1962, mainstreaming it into communication and information science circles. His DOI theory has been widely accepted to explain how people in any given social system—organization or society—adopt a new idea, practice, or philosophy. Rogers proposed that four aspects, namely 1) the importance of the innovation itself, 2) communication channels, 3) time, and 4) the social system, affect the adoption process in five stages. The first stage occurs when early innovators of a new idea spread the word so that more and more people become aware of the idea, even though exposure might lack complete information about the innovation. Next, interested people are persuaded by peers and seek additional information. In the third stage, people apply the innovation to their present or future life, and then decide whether to try the idea or novel practice. Stage four is characterized by people implementing the innovation into their life, and the final stage is reached when people decide to fully use the innovation (Kaminski, 2011, pp. 4-5). Over time, the idea becomes diffused among the system until it reaches critical mass. In 1943, Bryce Ryan and Neal Gross introduced five widely familiar categories of innovation adopters: innovators, early adopters, early majority, late majority, and laggards. Non-adopters are sometimes considered a sixth category (Kaminski, 2011).

Rogers (1995, pp. 15-16) warned that not all innovations are equivalent units of analysis, explaining that some ideas may take only a few years to reach widespread adoption and others require decades. He further clarified that multifarious characteristics could explain various adoption rates. Innovators should consider the a) *relative advantage* of the innovation over a previous way of doing something, measured in terms of economics, social capital, convenience, and satisfaction; b) *compatibility* of the innovation with existing values, past experiences, and needs of potential adopters; c) *complexity* of understanding or using the innovation; d) the degree of *trialability* for people to experiment with the innovation; and e) *observability* of people using the innovation.

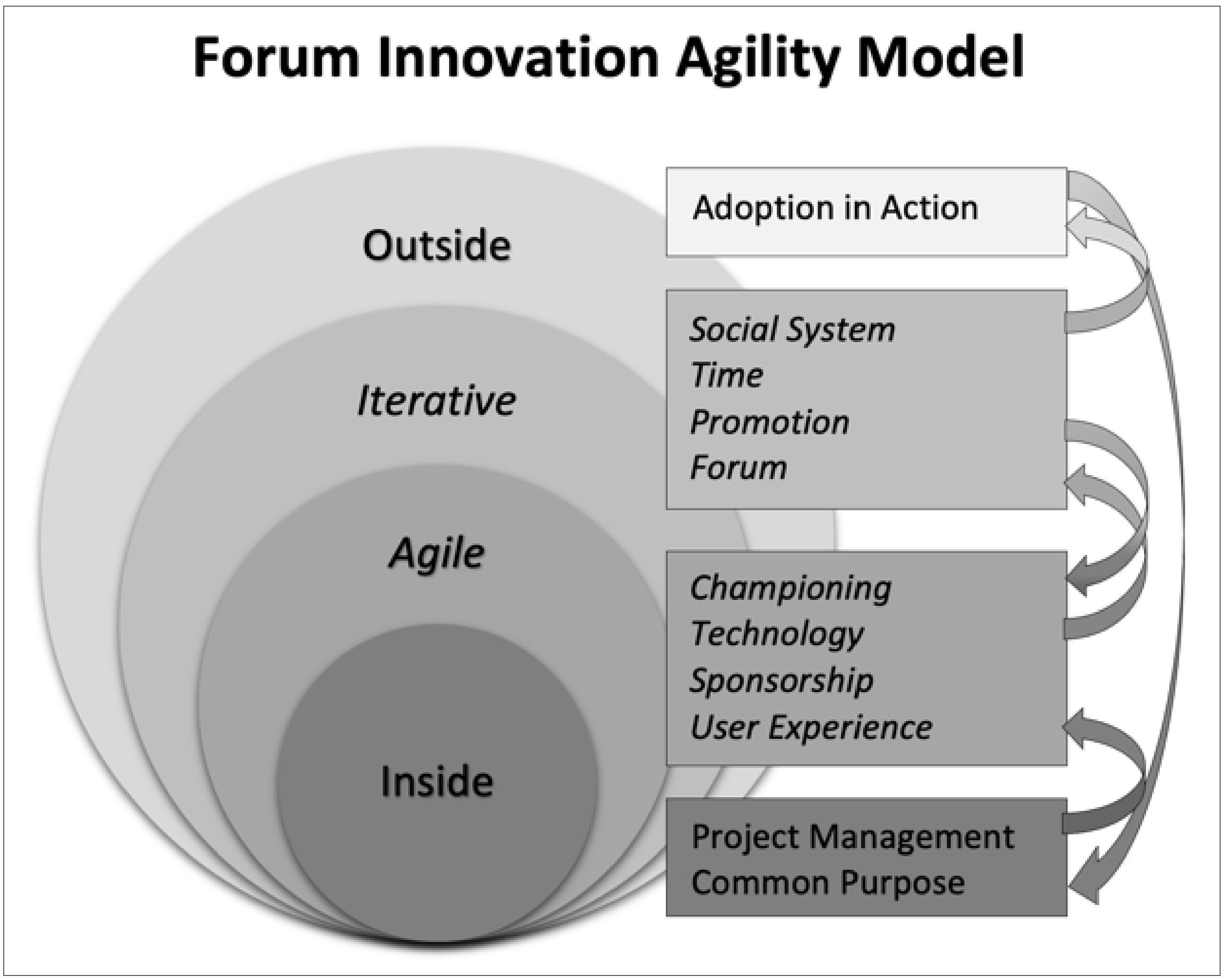
For example, the adoption of an incompatible innovation (e.g., transitioning from traditional pre-publication peer-review to open post-publication peer review) suggests a slow process of diffusion. Simpler to understand innovations (i.e., the ability to post a comment on a webpage) are adopted faster, and the likelihood of observing the results of adopting an innovation (i.e., seeing an increasing number of people commenting on webpages) stimulates peer discussion and encourages potential adopters to request more information about the innovation (i.e., how can I post a comment on a webpage?).

5.2 Forum Innovation Agility Model

Participation in an online forum can be simultaneously influenced by technology acceptance and community factors. Innovations involving ICTs are complex—with solutions initially unknown and requirements that will likely change (Rigby, Sutherland, & Takeuchi, 2015, para. 15). PubMed Commons was typical of an ICT innovation. Eleven confounding factors affected its use and are summarized in the forum innovation agility model (see 5.1). This socio-ecological model is a) based on Rogers' DOI theory, b) informed by my explanatory case study of PubMed Commons, and c) inspired by "That's Outside My Boat: Letting Go of What You Can't Control" (Jones & Doren, 2013). In his book, NBC sportscaster Charlie Jones shared advice from Olympic rowers. When asked about distractions that would cost them a medal, most athletes gave the same answer to his question: "That's outside my boat." While outside circumstances (e.g., wind, waves) could alter what rowers did inside their boat, they had to focus on things they could control.

Figure 5.1

Forum Innovation Agility Model



Factors affecting the PubMed Commons forum are organized in four layers on a continuum ranging from inside to outside the control of the forum's innovators and organizers. The model provides organizers with a list of factors they should consider, ensuring their focus remains on what they can control. Circumstances outside their control should be iteratively monitored using a variation of actions based on Rogers' four elements of diffusion. Findings might require agile adjustments to ensure forum adoption and sustainable use. Intended and unintended consequences of the adoption in action should be evaluated to determine if a new common purpose is warranted and changes to agile factors should be made.

Factors 1 (stating a common purpose) and 2 (project management) are completely under the control of this group, thus, figuratively, inside their boat (i.e., "Inside" layer of factors in the model depicted in Figure 5.1).

"Agile innovation methods have revolutionized information technology" (Rigby, Sutherland, & Takeuchi, 2016, para. 1). Agile factors (i.e., features that can pivot quickly) are created by project teams that plan in detail only forum functionality that won't change before execution. The team then starts developing contingency plans to adjust to changes outside their control (para. 12). Factors in the "Agile" layer include 3) user experience, 4) sponsorship, 5) technology, and 6) championing. Measuring outside forces should be done iteratively (i.e., repeatedly) to signal when contingency plans should be enacted. This means that Agile and Iterative factors consistently influence one another. Factors in the "Iterative" layer are defined by Rogers' four main elements in the diffusion process: 7) social system, 8) promotion, 9) the innovation itself (in this case study—the forum network), and 10) time (Rogers, 1995, p. 10-24). Innovators and organizers should eliminate distractions (i.e., factors) that are "Outside" their control and focus on results of their stated purpose. Factor 11 addresses this and suggests that organizers should evaluate how the innovation is adopted in action. Careful examination will indicate necessary changes in the agile and iterative processes to ensure successful adoption.

5.2.1 Inside Layer

In Jones' book, Olympic rowers expressed that they were solely interested in what they could control to win a medal. Inexperienced rowers are surprised by the instability of their boat's long axis and narrow, semi-circular hull, which causes their boat to roll from side to side. A rolling boat is difficult to row, thus slow. The crew must learn to balance the boat. Although a rowing boat has a *tendency* to roll, the rolling of an unbalanced boat is caused *entirely* by the crew (The Rowing Club, 2013). When introducing a new online forum, its use could be affected by many things. Innovators are completely in control of declaring a common purpose and managing the project. The first two factors affecting the adoption and use of PubMed Commons—common purpose and project management—deserved the utmost focus and attention.

Common Purpose. Rogers (1995) maintained that communication about innovation is a main element in the diffusion of a new idea, practice, or philosophy. Exploiting channels of communication has the potential to mitigate perceived complexities about the innovation and facilitate a better understanding among adopters (pp. 17, 18). One advantage of this explanatory case study research is that as an epistemological outside researcher, I experienced the roll-out of the forum as a potential adopter. Initial messages about PubMed Commons' purpose shared by its apparent organizers (e.g., NIH, NCBI, NLM, PubMed Commons spokespersons, and scientists who were among the forum's innovators and team leads) used a variety of similar, yet vague terms to communicate the purpose of PubMed Commons in articles, on their websites, and in blog and social media posts. These included "open discourse," "constructive criticism," "culture of commenting," "high-quality discussions," "debates," "sharing of information," and "social commenting" for promoting "collaboration among the scientific community."

For example, an Oct. 22, 2013, announcement on the NCBI Insights (2013a) blog stated,

> NCBI has released a pilot version of a new service in PubMed that allows researchers to post comments on individual PubMed abstracts. Called PubMed Commons, this service is an initiative of the NIH leadership in response to repeated requests by the scientific community for such a forum to be part of PubMed. We hope that PubMed Commons will leverage the social power of the internet to encourage constructive criticism and high-quality discussions of scientific issues that will both enhance understanding and provide new avenues of collaboration within the community. (para. 1)

Tibshirani (2013)—Stanford METRICS Lab professor, PubMed Commons innovator/scientist team lead, and first commenter on the forum—shared an Oct. 22, 2013, announcement about PubMed Commons on his website, reasserting the need for a commenting system.

> We all read a lot of papers and often have useful things to say about them, but there is no systematic way to do this—lots of journals have commenting systems, but they're clunky, and, most importantly, they're scattered across thousands of sites. Journals don't encourage critical comments from readers, and letters to the editor are difficult to publish and given too little space. If we're ever going to develop a culture of commenting on the literature, we need to have a simple and centralized way of doing it. (para. 1)

PubMed Commons lead editor Bastian (2013, para. 8) wrote a *Scientific American* blog post describing PubMed Commons as an "opportunity to draw together debates about a publication in a frequently-visited central research hub." Two weeks later, the "PubMed Commons team" posted on the new PubMed Commons Blog (see Nov. 26, 2013, in Appendix A) that the forum was:

> a pilot system that enables authors' discussion and sharing of information about publications in PubMed. Exploring options to ensure a vibrant and useful forum for discussion of scientific publications will be a key focus on the next stage of the pilot. (paras. 1, 6)

NIH director Collins and principal deputy director Tabak (2014) used similar positive commenting-culture language in a Jan. 30, 2014, article in *Nature,* reiterating the forum's purpose for "open discourse about published articles [where] authors can join and rate or contribute comments" (p. 3). Collins (2014) restated this sense of comradery in his Aug. 2014, NIH Director's Blog, comparing science to a team sport that deserved "as many avenues as possible through which to interact" (para. 1). He encouraged authors to participate in online discussions, update and receive feedback on their papers from scientists around the globe, and link to datasets and non-biomedical journals that might be overlooked.

Stanford professor and PubMed Commons' scientist team lead Steven Goodman (2014) answered the question "what is PubMed Commons?" in an Apr. 22, 2014, videotaped message on YouTube by describing it as a

> social commenting function that you see almost everywhere else on the web… whereby scientists could comment on the published work that appeared in PubMed, just like when you go on Amazon you can see how many reviews there have been done. (0:50)

He described having no reviews in the PubMed indexing catalog as a critical void.

One noticeable absence in messages describing PubMed Commons was the use of the term "post-publication peer review (PPPR)." PPPR had already become a hotly debated topic in the scientific community. The term PPPR was first mentioned in the PubMed Commons Blog on Dec. 17, 2014, and was used only twice more on Apr. 6, 2017, and May 28, 2017 (see Appendix A). However, social media communication from early adopters and behind-the scenes forum innovators like PLoS co-founder Michael Eisen (2013) implored NIH grantees who want to see science communication improve to see PubMed Commons as "great opportunity for us to make PPPR real" (para. 7).

It wasn't until a year after the open pilot launch of PubMed Commons that the term "PPPR" began creeping into the official vernacular. In her *PLoS Medicine* article entitled "A Stronger Post-Publication Culture is Needed for Better Science," Bastian (2014) wrote that PubMed Commons enables "post-publication commenting." She explained—highlighting her comments were her own and did not reflect those of the NIH—that while commenting overlaps with PPPR, it "does not encompass all of that activity" (p. 1). She affirmed "the negative 'yin' of criticism, correction, retraction, and failed replication" inherent in post-publication commenting and contrasted it with the "positive 'yang' aspect [of] incorporating research aftercare" (p. 1).

Another baffling oversight in describing the purpose for PubMed Commons was organizers using two terms—"discussion" and "discourse"—to describe posting activity. Among my colleagues, we discussed that each communication act would manifest in a different type of forum, warranting separate measures for evaluating adoption. While a discussion implies an "activity in which people talk about something and tell each other their ideas or opinions" (Cambridge University Press, 2021a), discourse is understood as a speech or piece of writing about a particular, usually serious, subject (Cambridge University Press, 2021b). In short, the former implies a social exchange; the latter is an independent statement. When asked about distinguishing between a social network of commenters and a *network of independents,* Bastian explained that many comments do not require a response, thus the reason why the forum had so many single-comment posts (H. Bastian, personal communication, October 28, 2016). She said PubMed Commons was not a social network and the aim of the forum wasn't the conversation, admitting that it was unusual for back-and-forth postings. She went on to say that "PubMed Commons never needed to be huge, just open." She asked rhetorically, "for

1,000,000 articles a year, should there be 50% receiving comments?” and answered emphatically, “No. That would be a massive amount of information, probably most of it not worth reading.” Bastian further explained that the purpose of PubMed Commons was addressing the fact that “our government is funding research, and the NIH wants transparency.” Admittedly, these words quelled my fears about the longevity of PubMed Commons, given I and others had noted the low adoption rate (Lane, 2016a; Lane, 2016b; Ramos, Davies, Grant, McMinn, Nunn, & Wilson, 2016).

Clearly, for early forum adopters there were obvious mixed messages about purpose. Were commenters expected to engage in a social discussion or post a scholarly communication statement? Did organizers intended for there to be “scientist-to-scientist interactions” (Collins & Tabak, 2014, para. 12) or should commenters “offer peer review after publication” (Goodman, 2013)? Along those lines, who qualified as a “peer” on PubMed Commons? For example, does authoring an article indexed on PubMed make you a peer regardless of experience or subject knowledge? According to United States Geological Survey (USGS) Associate Director for Administration Jose Aragon (2016), a qualified peer possesses the appropriate education or expertise to offer comments on the work of others in the same field (para. B). Could a cardiologist comment on vaccine research? Similarly, a community buzz on social media was questioning whether junior researchers would challenge seasoned ones about methodology and conclusions.

This loss of message control was an unintended consequence of a large group of organizers and impassioned early adopters who were using the internet—the very channel PubMed organizers intended to improve communication among scientists— to simultaneously innovate a large-scale commenting system and make a cultural shift in scholarly communication. As a result, messages about PubMed Commons failed to facilitate a better understanding of the innovation among adopters, as misgivings about PubMed Commons spread among colleagues on social media. Sharing the news about the forum was often overshadowed by personal opinions about PPPR.

For example, on Oct. 23, 2013, Carl Heneghan, a PubMed Commons beta tester invited by Bastian, shared on his personal blog that it was impossible for him not to think the forum represents a "significant moment in post peer review commentary" (para. 2). A model early adopter, Heneghan provided screenshots (an example of Rogers' observability) and shared his experience that the forum was "pretty simple to use" (para. 4, 5). He also linked to FAQs on the PubMed Commons homepage, the PubMed Commons Twitter page, and a *Scientific American* blog post written by Bastian. His final thoughts were "this is going to be massive" (para. 10).

In a Dec. 20, 2013, *SciELO in Perspective* blog post, Lillian Nassi-Calo, Coordinator of Scientific Communication at BIREME/PAHO/WHO wrote that "recent advances in scientific communication worldwide advocate the publication of open peer reviews, the use of social networks and altmetrics to promote the dissemination and discussion of research results" (para. 1). She introduced PubMed Commons to her readers as "a system that enables scientists to post open comments about scientific papers" (para. 1). The *Journal of the Canadian Dental Association* (2013) introduced PubMed Commons as an open peer review system where "readers of scientific literature (not journal editors) drive the discussions" (para. 2), adding that authors are using PubMed Commons to post corrections, change their conclusions or update people on their work" (para. 2).

An editorial for *the Journal of Histochemistry & Cytochemistry* written by John Couchman (2013) from the University of Copenhagen picked up on the term "peer" and immediately wondered if the PubMed Commons "commentary is 'peer' reviewed," asking his audience "will the comments be applied to papers by contributors with real expertise in the area?" (p. 9). In an Oct. 22, 2013, post highlighting the pros and cons of re-reviewing published papers, Aimee Swartz framed peer review on Daily News' *The Scientist*, characterizing it as "what irks [scientists] most about publishing" (para. 1). She explained that "the process has been blamed for everything from slowing down the communication of new discoveries to introducing woeful biases to the literature," adding "few believe peer review is capable of accomplishing what it purports to do—ensuring the quality of published science" (para. 1). Swartz mentioned the inconvenience of commenting journal by journal, communicating that the new "post-publication peer review system housed on the oft-accessed NCBI biomedical database" would be "simple" (para. 10).

Some scientists disputed such rave reviews, predicting the forum would fail. On his website, Pawel Niewiadomski (Oct. 29, 2013) said "I am willing to bet good money that it will not succeed at what its main goal is, namely enabling open and objective discussion of the merits of published papers" (para. 1). He continued, "it can hurt you in a major way if you step on someone's toes by leaving a critical comment on their paper" (para. 3).

Recall from Chapter 4 that articles receiving the most comments on PubMed Commons mirrored this intertwined purpose of introducing a forum for scientific discussion and debate over aspects of changes in scholarly communication.

Perhaps more damaging than this duality in purpose was that the key focus of the next phase of forum introduction explicitly stated by the PubMed Commons team (2013a) was "hosting authors' discussion and sharing information about publications in a vibrant and useful forum" (para. 1). Bastian's negative "yin" of "criticism, correction, retraction, and failed replication" began to take center stage on her and Vaught's agenda. For example, the only peer-review article published by the co-editors during their time managing PubMed Commons was "Concern Noted: A Descriptive Study of Editorial Expressions of Concern in PubMed and PubMed Central" (Vaught, Jordan, and Bastian, 2017), which appeared in BioMed Central's *Research Integrity and Peer Review.* The PubMed Commons Team also uploaded data they had collected about "evaluations of post-publication activities at PubMed and associated databases, [including] letters to the editor, PubMed Commons comments, corrections, findings of the HHS Office of Research Integrity, editorial expressions of concern, and retractions and withdrawn publications" on an Open Science Framework (OSF) "PubMed Commons Post-Publication Project" wiki page (last updated Sept. 23, 2019).

On Mar. 21, 2017, in a videotaped presentation entitled "Post-publication Peer Review and Certificate Systems," which is archived on NCBI's YouTube Channel, Bastian (2017) admitted that

> [PubMed Commons is] not a place that people really discuss things. They discuss things elsewhere, so it ends up being about four publications a day that are getting, being commented on. And so you have quite a wide range of commenters, quite a wide range of types of articles being commented on, but not necessarily a lot of commenting, which is not—we don't think—is a bad thing. (21:15)

She went on to detail a new focus she and her team had for the forum:

> But now we're kind of working and taking a real perspective on post-publication activity overall in PubMed and what are all of the kinds of things. We've had a history of people tending to study things like retractions and findings of research

> fraud and so on, which are actually a miniscule part of post-publication activity—it's like less than ½ percent of what's in PubMed. The real kind of bulk of what's happening is the things like, oh… letters to the editor, erratum, and those kinds of things. You see, that's the really big area, and we are starting to go through working on analyzing what these are and what the implications of those are and being able to put a lot of this in perspective and find ways we can develop tools that address some of those things I talked about earlier—what could actually get this into people's workflows and those kinds of things. That's what we're really looking at. How can we do it better? One of the first things we're doing is tagging editorial expressions of concern in PubMed, which will be coming soon. Here you see a rise in this. It's still a very unusual thing to do, but there is a real rise in the use of this mechanism by journals. (23:52)

After watching the video, I wondered what about building a vibrant and useful forum? No official communication about this expanded mission or pivot in purpose was shared on the NCBI Insights blog, the PubMed Commons Blog, the NIH Director's Blog, or any other outlet managed by PubMed Commons organizers.

Rogers (1995) warned that an inconsistent message of purpose weakens the essence of the diffusion process to have one individual communicate a new idea to one or several others through an information exchange. He reported diffusion research that showed most people depended not on formal scholarly communication about an innovation, but rather on personal assessments "conveyed to them from other individuals like themselves who have previously adopted the innovation" (p. 18). In this way, posting and comments shared on social media and in journal forums were critically important for communicating the purpose and value of PubMed Commons. Again, my review of these types of personal messages showed that adopters believed the purpose and their role on the forum was to solely and openly focus on and engage in PPPR.

5.2.1.1 Project Management

"If you fail to plan, you are planning to fail" goes the famous quote attributed to Benjamin Franklin (Avwontom, 2016). Rogers (1995) cautions individuals implementing a new idea to not become complacent believing that an innovation will take off on its own merit. "Most innovations, in fact, diffuse at a disappointingly slow rate" (p. 7). In other words, don't be hoodwinked into believing if you build it, they will come. Innovation launch and adoption require work.

Information about behind-the-scenes project management for PubMed Commons was not widely available. Project innovation and development history is available from the *Stanford Medicine News Center* (Spector, 2013). The following brief summary provides necessary context. In Dec. 2011, Tibshirani conceptualized a place where scientists could exchange ideas, ask questions about methods/techniques, offer suggestions, make comments on each other's work, and even offer praise (para. 7). Think of this venture as somewhat of a positive, asynchronous, virtually hosted meeting of peers who part of the larger invisible college. Authors could also update their work or post corrections. Tibshirani approached Stanford colleague Pat Brown—co-founder of the Public Library of Science (PLoS) open-access publishing enterprise, who had introduced and closed a comparable platform due to low adoption rate. A videoconference was held with NCBI director Lipman, who told them that although he had been thinking about something similar for over a decade, he was concerned how his organization could manage it. He added that until recently he didn't think the idea would get NIH leadership support. Lipman said he liked that the request for the platform was coming from the scientific community, not within NIH (Spector, 2013, para. 15, 16). Brown's friend and former director of the National Cancer Institute, Harold Varmus, affirmed the forum idea.

In Sept. 2012, Lipman proposed the platform to the NIH steering committee, composed of NIH director Collins and 10 directors from NIH institutes and centers. By that time, he had decided that Tibshirani and Brown should be project leads, getting others involved, and fully developing the idea. Lipman wanted to see if there was an active group of scientists who cared enough about the idea. Tibshirani served as organizer and worked with Brown to form a 300-member community to set guidelines and commence commenting on a beta version of the forum (para. 18-20). Site design was simple, requiring minimal NCBI involvement; an unmoderated forum model would keep staffing negligeable. Notably, during the development phase, up to 20% of the community voiced concerns about commenting repercussions or just didn't see the point (para. 33).

PubMed Commons was envisioned to eventually allow nonscientists (e.g., biomedical journalists and patients with personal expertise) to comment alongside scientists. Posting anonymity was the most contentious issue (para. 21).

The most visible communication from NCBI about strategic project planning with a schedule came in a Nov. 26, 2013, post on the PubMed Commons Blog. Three objectives were identified: 1) uptake and reputation, 2) quality and impact of comments and discussion, and 3) sustainability.

> We are establishing a working group to advise us during the next stages of the pilot and its evaluation. And we look forward to community discussion, too. The upcoming release marks the start of evaluation of PubMed Commons. Evaluation results will be considered at 3 months and 6 months, with the final report on the pilot anticipated after 9 months. Three key areas will be our focus: uptake and reputation, quality and impact of comments and discussion, and sustainability. We will be blogging more about

> comments being made in the Commons, what we're learning, and explaining more about aspects of the system. We will be trying out a Twitter chat too, so keep your eye out on @PubMedCommons for the announcement. (The PubMed Commons team)

I found no evidence of a community discussion or milestone reports. A follow-up post about the evaluation came Dec. 17, 2015, on the PubMed Commons Blog. I was unable to locate the presaged final report.

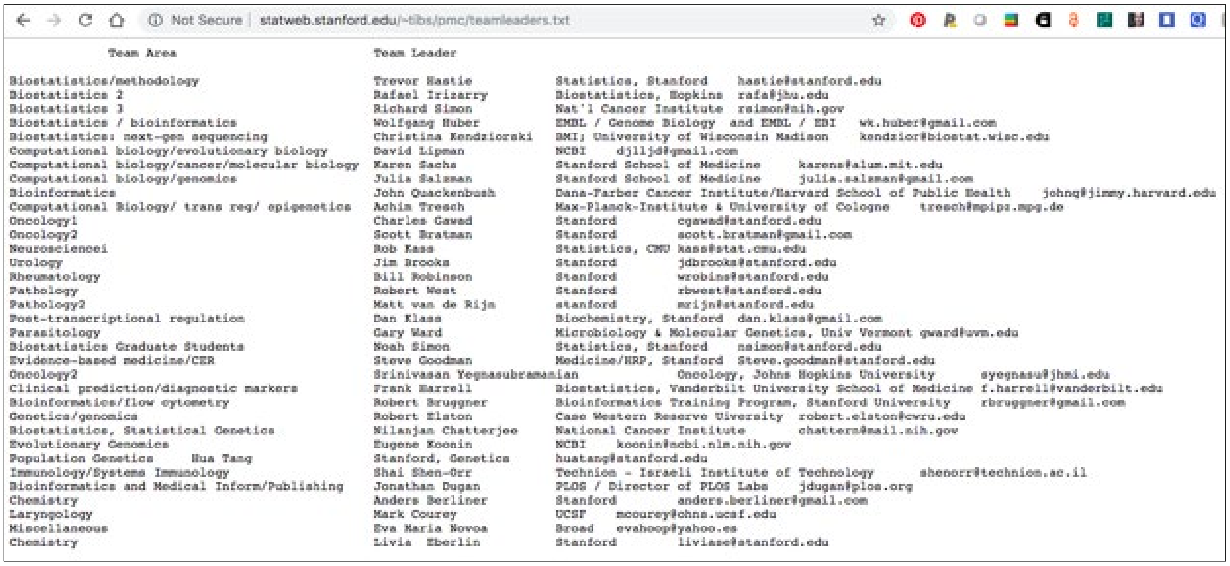
> We are pleased to announce that PubMed Commons is here to stay! After developing and piloting the core commenting system for PubMed, a pilot of journal clubs was added. And we have completed a major internal evaluation of the use of the Commons. We aim to publish that soon, so stay tuned to this blog or Twitter for news on that. (The PubMed Commons team)

Again, I was never able to find communication about the major internal evaluation. The limited number of posts on the NCBI Insights blog (seven) and PubMed Commons Blog (eight) that were made during the PubMed Commons project lifecycle gave few clues about project methodology or project management. Recall from the Sponsorship factor discussion that NCBI Director Lipman said, "this [PubMed Commons] is not an NIH project, except that NIH is allowing them to do it. We need to get the drive and direction from people who are going to be active in it" (Oransky, 2013, para. 23). This would imply that the strategic plan and buy-in for the project was the responsibility of the scientists who were innovating the forum and who would be using it (Oransky, 2013, para. 23). The information confirms the R&D history on the Stanford website.

The most tangible evidence of an organizing group was a document linked from Tibshirani's Oct. 22, 2013, announcement about the forum (para. 3), which is shown in Figure 5.2. Thirty-four "team leads" from diverse institutions with a variety of subject area expertise are identified.

Figure 5.2

PubMed Commons Team Leads and Areas of Expertise



```
Not Secure | statweb.stanford.edu/~tibs/pmc/teamleaders.txt

            Team Area                                   Team Leader

Biostatistics/methodology                         Trevor Hastie          Statistics, Stanford    hastie@stanford.edu
Biostatistics 2                                   Rafael Irizarry        Biostatistics, Hopkins  rafa@jhu.edu
Biostatistics 3                                   Richard Simon          Nat'l Cancer Institute  rsimon@nih.gov
Biostatistics / bioinformatics                    Wolfgang Huber         EMBL / Genome Biology  and EMBL / EBI    wk.huber@gmail.com
Biostatistics: next-gen sequencing                Christina Kendziorski  BMI; University of Wisconsin Madison    kendzior@biostat.wisc.edu
Computational biology/evolutionary biology        David Lipman           NCBI    djlljd@gmail.com
Computational biology/cancer/molecular biology    Karen Sachs            Stanford School of Medicine     karens@alum.mit.edu
Computational biology/genomics                    Julia Salzman          Stanford School of Medicine     julia.salzman@gmail.com
Bioinformatics                                    John Quackenbush       Dana-Farber Cancer Institute/Harvard School of Public Health    johnq@jimmy.harvard.edu
Computational Biology/ trans reg/ epigenetics     Achim Tresch           Max-Planck-Institute & University of Cologne    tresch@mpipz.mpg.de
Oncology1                                         Charles Gawad          Stanford        cgawad@stanford.edu
Oncology2                                         Scott Bratman          Stanford        scott.bratman@gmail.com
Neurosciences1                                    Rob Kass               Statistics, CMU kass@stat.cmu.edu
Urology                                           Jim Brooks             Stanford        jdbrooks@stanford.edu
Rheumatology                                      Bill Robinson          Stanford        wrobins@stanford.edu
Pathology                                         Robert West            Stanford        rbwest@stanford.edu
Pathology2                                        Matt van de Rijn       stanford        mrijn@stanford.edu
Post-transcriptional regulation                   Dan Klass              Biochemistry, Stanford  dan.klass@gmail.com
Parasitology                                      Gary Ward              Microbiology & Molecular Genetics, Univ Vermont gward@uvm.edu
Biostatistics Graduate Students                   Noah Simon             Statistics, Stanford    nsimon@stanford.edu
Evidence-based medicine/CER                       Steve Goodman          Medicine/HRP, Stanford  Steve.goodman@stanford.edu
Oncology2                                         Srinivasan Yegnasubramanian            Oncology, Johns Hopkins University      syegnasu@jhmi.edu
Clinical prediction/diagnostic markers            Frank Harrell          Biostatistics, Vanderbilt University School of Medicine f.harrell@vanderbilt.edu
Bioinformatics/flow cytometry                     Robert Bruggner        Bioinformatics Training Program, Stanford University    rbruggner@gmail.com
Genetics/genomics                                 Robert Elston          Case Western Reserve Uiversity  robert.elston@cwru.edu
Biostatistics, Statistical Genetics               Nilanjan Chatterjee    National Cancer Institute       chattern@mail.nih.gov
Evolutionary Genomics                             Eugene Koonin          NCBI    koonin@ncbi.nlm.nih.gov
Population Genetics     Hua Tang                  Stanford, Genetics     huatang@stanford.edu
Immunology/Systems Immunology                     Shai Shen-Orr          Technion - Israeli Institute of Technology      shenorr@technion.ac.il
Bioinformatics and Medical Inform/Publishing      Jonathan Dugan         PLOS / Director of PLOS Labs    jdugan@plos.org
Chemistry                                         Anders Berliner        Stanford        anders.berliner@gmail.com
Laryngology                                       Mark Courey            UCSF    mcourey@ohns.ucsf.edu
Miscellaneous                                     Eva Maria Novoa        Broad   evahoop@yahoo.es
Chemistry                                         Livia  Eberlin         Stanford        liviase@stanford.edu
```

Data analysis of the commons_archive.csv file showed that only nine of the 34 team leads commented on PubMed Commons. See Table 5.1. I reiterate, with all good intentions, something happened on the way to the forum.

Table 5.1

PubMed Commons Team Leads Who Commented on PubMed Commons

Team Lead	Posting Dates	Total
Brooks, Jim	06-19-13	1
Dugan, Jonathon	06-18-13	1
Huber, Wolfgang	06-15-13, 06-16-13, 07-02-13, 01-30-16, 04-12-16	5
Koonin, Eugene	02-24-14	1
Quackenbush, John	06-16-13, 09-30-15, 04-18-16	3
Salzman, Julia	06-13-13	1
Simon, Noah	06-18-13	1
Ward, Gary	06-24-13, 10-20-13, 11-17-13, 11-20-13, 12-06-13, 12-10-13, 11-04-14, 11-09-14	8
West, Robert	01-28-17	1

According to Joseph Lukas (2014) from the Project Management Institute (PMI), there are five key items required for all projects: 1) project charter, including objectives and success criteria; 2) project requirements; 3) work breakdown structure (WBS); 4) the plan, including things like milestone gates, quality checks, risk management/mitigation plans, time-phased budget, resource allocation; precise scheduling; and, finally, 5) communications. A project charter is “a document issued by the project initiator or sponsor that formally recognizes the existence of a project” (PMI, as cited by Lukas, para. 9). Project requirements describe what is needed to achieve objects and serve as the foundation upon which project scope and plan are developed. Lukas shared that the leading problem during project development is due to conflicting priorities or opinions (para. 13). WBS defines the entire project scope and its tangible deliverables, which aids in determining processes for the plan. It also addresses status and progress reports and a baseline for change management. “A poorly constructed or incomplete WBS results in scope creep, unclear work assignments, schedule dates slippage, and cost overruns” (para. 15). Deliverables are project specific; the plan should be flexible to accommodate processes and needs to ensure project success. Sometimes a simple spreadsheet is sufficient for a project, and others will require cost estimates and budgets. Excellent communications are common among successful projects. Determining communication needs requires deciding a timeline specifying what information to share with whom using which processes and media.

A critical understanding for project management is that the project lifecycle is iterative and requires agility. Agile does not mean planning is not done. Rather, it means there is a project strategy, roadmap, release plan, iteration plan, and daily plan (para. 20). A project methodology like the one detailed in the previous paragraphs leads to improved project results. Of critical importance is an understanding that "Across all industries, organization types, and project sizes, success or failure is largely dependent on having the right levels of team commitment, stakeholder buy-in, and support" (Aziz, 2014, para. 1).

Discussions earlier in this chapter suggesting weaknesses in the PubMed Commons project's statement of purpose, sponsorship, and technology factors could have affected its adoption and success. It is difficult to make this determination, however, due to a) insufficient project management external communications, especially regarding measures for successful forum implementation and adoption. Recall Bastian's statement that "slowly building the case for the value of commenting is part of changing the perception of how you measure success, and with regard to PubMed Commons, the number of comments isn't the answer" (H. Bastian, personal communication, October 28, 2016).

Considering the extent of project management processes outlined above and after reviewing data and analyses presented in Chapter 4, one might wonder whether the PubMed Commons innovation and implementation followed a project methodology or utilized a project manager. If these were invoked, would the forum's outcome have been different? Was there sufficient sponsorship investment and buy-in from a committed social system of supporters?

Without sponsorship support and project team champions, project managers are powerless, and successful project completion is uncertain. Furthermore, "Even when the scientific elements of project management are thoroughly developed and applied, the risk of project failure is imminent if 360° stakeholder buy-in is inadequate, or fluctuates throughout the project" (Aziz, 2014, para. 1).

I maintain that findings reported in Chapter 4—number of commenters, level of participation, and rate of adoption during the five-years the forum was operating—suggest insufficient project management. The forum innovation agility model suggests that had the organizers been aware of forces happening outside their control (i.e., agile and iterative layer factors), they could have enacted the project's formal contingency plans.

5.2.2 Agile Layers

5.2.2.1 User Experience

"User experience focuses on having a deep understanding of users, what they need, what they value, their abilities, and also their limitations'' (Usabililty.gov, 2021, para. 1). Having users find value in what is being provided to them is key. Of consideration are desirability, usability, findability, accessibility, and credibility (para. 3), which are commonly expressed in online forum guidelines. These rules of engagement are a logical follow-up to "Inside" factors (e.g., common purpose and project management), although they must be informed by users, who are outside the control of organizers. When devised in tandem with technology, features like interface design, interaction design, content strategy, and, again, accessibility (para. 5), can improve interaction and perceptions about a product or service.

At the most fundamental level, the adoption of PubMed Commons was influenced by the forum's "Guidelines" (see Appendix B). Like the statement of purpose, clearly communicating guidelines for usage helps adopters understand expected behaviors. "The objective of the site guidelines was to raise the bar for politeness. This is a complex thing" (H. Bastian, personal communication, October 28, 2016). When setting up the forum, organizers were concerned about protecting the reputation of researchers, thus “it was clear that guidelines were important.”

In the aforementioned 2017 videotaped presentation, Bastian briefly reviewed criteria under which the forum was developed:

> [Commenters] had to be an author of a publication in PubMed to be able to comment—no anonymous accounts or no pseudonyms; post comment moderation only; and with the exception of a quick filter (or automatic filter it goes through) and then you comment; and it is a creative commons license that you’ve agreed to. (20:10)

Commenters agreed to “disclose potential conflicts of interest” and “make comments that are directly relevant to the particular work in PubMed on which they are commenting.”

PubMed Commons guidelines also stated that comments should not include “allegations of misconduct on the part of authors, reviewers, editors and publishers” (Riegelman & Bakker, 2018, p. 42), which conflicted with an understood purpose to use post-publication commenting to self-correct science (Collins & Tabak, 2014). Commenter violations of this and an agreement to avoid “discriminatory, racist, offensive, inflammatory, unlawful, or derogatory language” became problematic, and personal testimonials by early adopters whose commenting privileges were revoked might have dissuaded adopters (see Figure 4.15). Recall that several forum users communicated that moderating was not transparent. Bastian confirmed this, explaining that comments were automatically posted, and then manually monitored by moderators. “The public can’t see if moderators block a post or if someone loses their posting privileges” (H. Bastian, personal communication, October 28, 2016). Quantitative content data in Chapter 4 reporting the number of comments deleted or edited by user and those removed by moderators evidenced the frequency of issues related to this part of the guidelines.

Despite the fact "comments should not have explicit commercial endorsements" (Oransky, 2013) and commenters should not "use PubMed Commons to spam or systematically promote a product, position or the members' own publications or to target others," this rule was visibly breached. For example, recall from Chapter 4 that Kath Wright's 111 comments on 111 articles were identical: "Other search filters are available from the InterTASC Information Specialists' Sub-Group Search Filter Resource at https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home." Hence, anyone aware of the guidelines and reading the comments would be led to believe that rules were not strictly enforced (i.e., there was wiggle room for interpretation). It is possible that the visibility of such violations could have affected forum adoption (Rogers, 1995).

Regarding registration, signing up didn't appear to be as simple as some early adopters indicated. Some social media posts reported that commenters had to be an NIH or Wellcome Trust grant recipient or get an invite from either a grant recipient or someone already participating in Pubmed Commons or belong to a group of at least 50 authors with email addresses ending .ac or .edu who wish to join en masse (PubMed Commons FAQ, 2013 cited in *Retraction Watch*, Oct. 22, 2013). After the closed beta testing of the forum (June 2013, to December 2013), potential adopters lamented on social media that they experienced difficulties becoming a commenter. Many shared with colleagues that after repeated issues to register, they gave up and did not adopt using the forum. This initial limited ability for early adopters to try to use the forum—in addition to requirement that commenters had to be authors of an article indexed on PubMed—grossly limited trialability, which Rogers (1995) indicated would make the adoption rate much slower (p. 16).

Similarly restricting for early adopters was an initial word limit for posting. "But commenters quickly got around this by adding several posts in succession" (M. Vaught, personal communication, October 28, 2016). This guideline was quickly revised.

Personal messages from early adopters who experienced some of difficulties with issues mentioned—those who found the forum's guidelines unfair/unenforced or who struggled with the registration process—likely aggravated perceived complexities and misunderstandings about the new forum. These communications might have deterred others from trying out the forum, which weakened the robustness of the diffusion process (Rogers, 1995, p. 17). Evidence of low adoption is clearly evidenced in data analyses reported in Chapter 4.

One of the most controversial (and distracting) conditions of commenting on PubMed Commons was associating a commenter's real name with their comment. Fervent opinions and reactions to this requirement sparked a great amount of communication about the forum among scientists on several channels (e.g., peer review literature, grey literature, social media), which could be considered either a bad or a good thing (Stanford GSB Staff). Nonetheless, the emotionally charged issue of anonymous vs. identified peer review challenges the chief tenet on which traditional scholarly communication among scientists has been grounded for nearly 200 years (Madden, 2000), and possibly affected the forum's adoption rate.

In her Oct. 22, 2013, "Post-publication Peer Review Mainstreamed" blog post, Swartz reported that "PubMed Commons chose to circumvent the downfalls of anonymity." She disclosed that NCBI Director (1989-2017) Lipman admitted there were "cogent and compelling arguments for both anonymous and identifiable commenting" (para. 17).

PubMed Commons innovator Tibshirani wrote in is Oct. 22, 2013, announcement about the forum that:

> One big issue that we have faced was the question of whether anonymous comments should be allowed. After much discussion, the group remained deeply split on this issue. Those wanting anonymous posts were concerned that many scientists, especially junior researchers, would be reluctant to make critical comments. But those opposed to anonymous comments believed that the quality of interchange would be higher if commenters were required to identify themselves. In the end, these differences weren't really resolved, and the decision was to start without anonymous comments and re-evaluate after the system had been fully public for a while. While debating this issue various proposals were put on the table for ways to allow participants to review and essentially sponsor the anonymous post of another participant. (para. 5)

Anonymous comments were never allowed on PubMed Commons, and the forum was not given enough time to evaluate its affects and implement contingency plans about this guideline.

Personal opinions about commenter identification were passionately debated on PubMed Commons, and as mentioned in Chapter 4, were part of threaded discussions on two of the most highly commented on article abstracts indexed in PubMed.

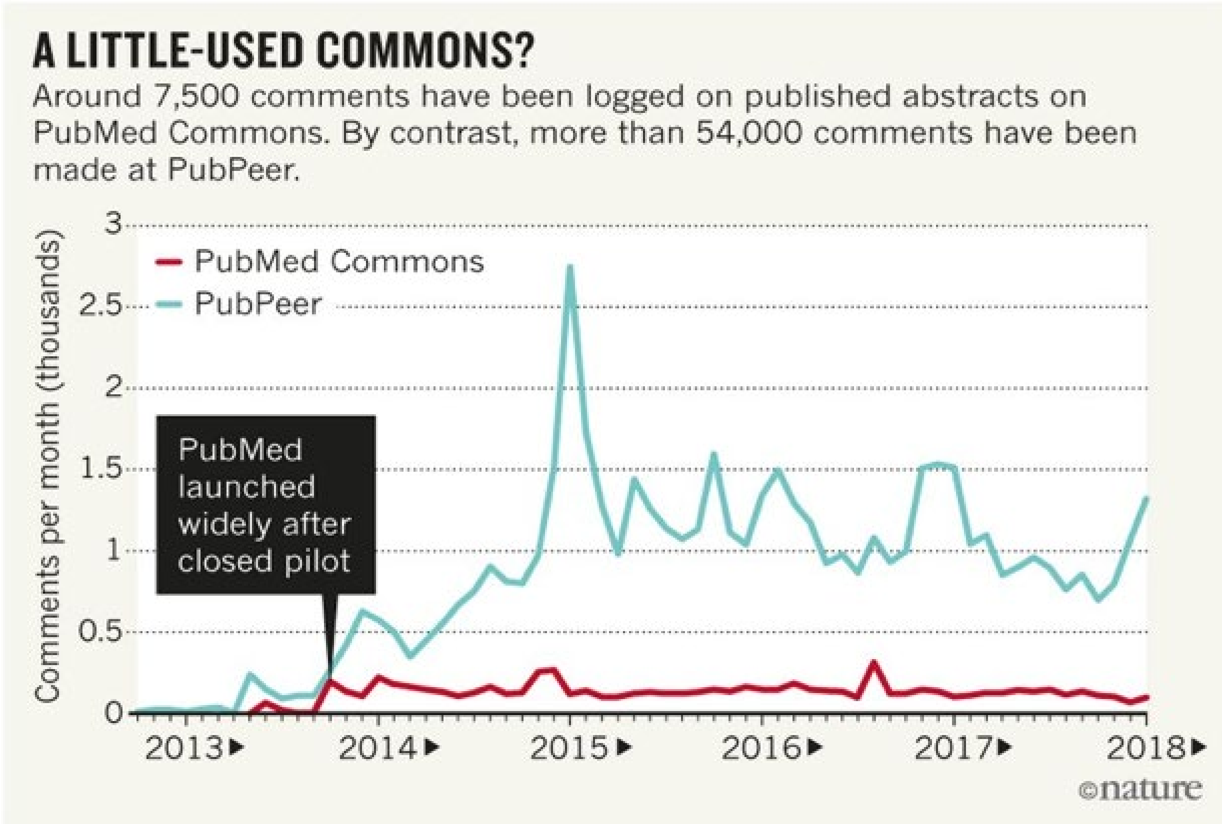
According to the articles' author,

> anonymity makes sense when reviews are offered in confidence to be assessed and moderated by an editor, someone whose identity is known and who takes responsibility for the decision informed by the reviews. Obviously, this same situation does not apply post-publication, not when the commenters enter into a discussion anonymously and the moderators are also unknown…. I concur with Hilda Bastian, who notes, on the one hand, the lack of reliable evidence to support the benefits of reviewer anonymity and, on the other, the importance of assessing whether commenters are 'outside their areas of expertise' [or] have conflicts of interest. Anonymity can conceal much mischief and do great damage. (Blatt, 2015, pp. 907-908)

He goes on to say that organizers at PubPeer—an alternative commenting system that allows anonymous comments—argue that anonymity and the possible low quality/bad faith comments are necessary to encourage frank and worthwhile discussion (p. 908).

Figure 5.3

"Nature" Figure Comparing Activity on PubMed Commons and PubPeer



PubMed Commons/Neil Saunders; PubPeer

Determining whether commenter identification positively or negatively affected the rate of PubMed Commons forum adoption requires more investigation, thus is suggested as future research. My environmental scan of literature comparing PubMed Commons with PubPeer shows that more users adopted PubPeer than PubMed Commons over nearly the same time period (see Figure 5.3). Even so, when comparing PubMed Commons with PubPeer, there are too many multifarious factors to conclude the higher adoption rate of PubPeer could be solely attributed to anonymous commenting.

Regarding any online behavior, this rule of engagement could significantly affect adoption rates independent of other guidelines. The issue of anonymous commenting is a societal one that touches human rights, free speech, trolls, misinformation, fake news, and open discourse (Rainie, Anderson, & Albright, 2017). The extent of these are well beyond the control of PubMed Commons organizers, although the ways in which anonymity affected the forum was a force that deserved attention. Results from an investigation could have prompted organizers to alter factors that were inside their control.

Sponsorship. Consider the concept of sponsorship as "the position or function of a person or group who vouches for, supports, advises, or helps fund another person or an organization or project" (Dictionary.com, 2021). Sponsorship can be an investment of time, talent, social capital, and/or financial support. The balance and emphasis of these can change at any given time. Unarguably, PubMed Commons' sponsorship was nothing but colossal. The forum was hosted on PubMed, had the blessing of NIH directors, and was stewarded by a community of widely recognized scientists who invested time, talent, knowledge from previous experiences, and the exponential power of their social capital to support the forum's start-up.

The presence PubMed has throughout the world is so very important when thinking about diffusing an ICT commenting vehicle innovation. Accessing and searching on PubMed is open (i.e., free), thus possible by anyone, anywhere who is connected to the Internet. "On an average working day approximately 2.5 million users from around the world access PubMed to perform about 3 million searches and 9 million page views" (Fiorunu, Lipman, & Lu, 2017, para. 2). PubMed's users (i.e., audience, therefore reach) amasses first world and third world medical and health care professionals, biomedical and life sciences scientists, and literally everyone with internet access: patients, families, friends, students, researchers, and so on.

PubMed is a service of the United States National Library of Medicine (NLM), which is part of the National Institutes of Health (NIH). The NLM is also the world's largest biomedical library. Its mission emphasizes the reasoning to sponsor PubMed Commons.

> The NLM pioneers new ways to make biomedical data and information more accessible and builds tools for better data management. … Leveraging its 184 [year]-history, NLM develops and applies innovative approaches to acquire, organize, curate, and deliver biomedical information across the United States and the world. NLM's advanced biomedical information services are among the most visited websites in the Federal Government. … NLM carries out its mission of enabling biomedical research, supporting health care and public health, and promoting health behavior by conducting research development on biomedical communications systems, methods, technologies, and networks and information dissemination and utilization among health professionals, patients, and the public. (NIH, 2020, para. 1, 5, 7).

The National Center for Biotechnology Information (NCBI) is a division of the NLM, established in 1988. Its employees are charged with "creating and maintaining over 40 databases for the medical and scientific communities as well as the general public. NCBI's core literature database is PubMed, which provides abstracts and citations for millions of articles from thousands of biomedical journals" (NIH, 2020, para. 38). PubMed Commons comments appeared below the abstracts on the PubMed GUI. "PubMed delivers a publicly available search interface for MEDLINE as well as other NLM resources, making it the premier source for biomedical literature and one of the most widely accessible resources in the world" (Williamson & Minter, 2017, p. 16).

PubMed Commons Innovators could not have selected a more visible, fitting sponsor than the NLM for the forum they envisioned. To reiterate, a) no database in the world has a greater presence or reach, b) none provides better accessibility to indexed records of biomedical literature, and c) the search engine has a massive, built-in, worldwide audience of users. Many PubMed Commons innovators and organizers agreed that PubMed was the perfect place for the post-publication commenting system that many scientists had expressed they wanted.

For example, forum innovator Eisen (2013) believed “The obvious place to build such a commenting/post-publication review system has always been directly in PubMed – it has everything, and everyone already uses it (para. 4). Stanford’s Goodman (2013) boasted PubMed’s breadth of reach, explaining that, “PubMed is the… giant catalog of scientific studies in all the top medical journals; actually, it goes well beyond the top—it’s many thousands of journals whose contents are indexed in PubMed. It’s like, a little bit like, a Google just for medical research” (0.01). BMC’s Cockerill hoped that PubMed’s reach could overcome after-publication commenting difficulties other venues experienced: “BioMed Central, like many other publishers, has long encouraged readers to comment on published journal articles, but has found that uptake of this functionality has been limited. PubMed’s exceptional breadth of content, combined with its huge traffic, could change that” (para. 2).

Perhaps the most significant advantage of hosting PubMed Commons on PubMed rather than on other seemingly obvious websites is that this placement reduced fragmentation of discussions. “Lots of journals have commenting systems, but they’re clunky, and, most importantly, they’re scattered across thousands of sites” (Tibshirani, 2013, para. 1). PubMed Commons enabled not only commenters, but also readers, to go to one resource, rather than spend time clicking from journal forum webpages to personal blogs to publisher websites, and so on.

PubMed Commons user Pedro Beltrao (2013), research group lead at European Molecular Biology Laboratory (EMBL)–European Bioinformatics Institute (EBI), agreed.

> Commenting systems have, for the most part, failed to work on the publisher's side and the hope is that this [PubMed Commons] might finally create a discussion forum with higher participation. The advantages here are a higher visibility and lower friction when compared with most publishers' existing commenting systems. (para. 3)

In light of these positives, recall that messaging from the PubMed Commons team brought into question, once again, the common purpose of the forum. These dimmed the significance of the widely embraced advantage that came from NIH, NCBI, and NLM sponsorship:

> The important thing to sort of understand is that the conversation about the comments and so happen elsewhere. Mostly, there's just one or sometimes two comments in PubMed Commons on a particular article. It's [PubMed Commons] not a place that people really discuss things. They discuss things elsewhere. (Bastian, 2017, 21:01)

What would be the benefit of the sweeping financial, talent, and time-commitment costs of the many federal government organizations and employees involved in the project if "open discourse about published articles [where] authors can join and rate or contribute comments" (Colins & Tabak, 2014 p. 3) wasn't taking place on PubMed Commons? This question was answered when PubMed Commons was discontinued.

While many worthwhile comments were made through the service during its 4 years of operations, NIH has decided that the low level of participation does not warrant continued investment in the project, particularly given the availability of other commenting venues. (NCBI Insights, 2018, para. 4)

Groups with a similar purpose (e.g., PubPeer and Retraction Watch) had also been attempting to build social systems, and traffic on their websites was growing (see Figure 5.3). Large sponsors like the John D. and Catherine T. MacArthur Foundation, the Helmsley Charitable Trust, and the Laura and John Arnold Foundation were supporting their work.

Perhaps more than the social presence granted by the U.S. government in hosting PubMed Commons on the PubMed GUI, was the time and attention given to the forum by a large group of (no doubt) busy scientists who had full time researching and, in some cases, teaching careers, among other responsibilities.

Given this obvious, grand scale investment, I was surprised to discover that NCBI contractors and PubMed Commons editors Bastian and Vaught were managing what appeared to be large PubMed Commons enterprise from a shared, small room in NLM building 38A, Lister Hill Center, 10th floor, office 1003N in Bethesda, MD. One of the driving forces behind my request for a meeting was to ask about the continuation of the forum, given its low adoption rate. I was looking for a sense of how big the operation was. As a PubMed user and outside observer/researcher of PubMed Commons, seeing the visible signs of the PubMed Commons logo, "Reader Comments" search filter, and trending comments on the PubMed GUI gave the impression of widespread support and significant financial and human capital investment from the NIH.

During the interview, I was disturbed to learn that the "PubMed Commons team," as they were known on the PubMed Commons Blog, seemed to have little staffing or technical support. When asked about data collection, Bastian said that the last time they had done any was during one quarter in 2014 when they looked at gender of commenters, amount of time between posts, and number of comments per person (H. Bastian, personal communication, October 28, 2016). Interested readers can reference Appendix G for a poster reporting results of the 2014-2016 data collection. Bastian admitted there should be additional data collection and analyses that compared comment content classified as criticism, curation, and non-specific. She was also interested in the type of publication being commented on.

When asked about future marketing for PubMed Commons, Bastian said that “PubMed Commons has a small group working on it, so they can only do so much.” She also said that the forum “doesn’t need to be expensive, it just needs to work.” At the end of our two-hour visit, Bastian concluded that “post-publication communication has value and should happen,” adding that part of what is needed is changing the perception of “how you measure success.” She pointed out that regarding PubMed Commons, “the number of comments isn’t the answer.”

My probing questions that day were sufficiently answered so that I left feeling somewhat reassured that the forum would continue, thus my investigation of the forum continued. My doubts persisted, however, about where and how the forum would get the support it obviously needed to grow its social system. Even so, the announcement of the forum’s discontinuation on Feb. 1, 2018, was a shocking blow. Data scientist Saunders (2018) summed up the disappointment I felt and that was expressed in many social media posts authored by individuals throughout the world: “It is a shame, in my opinion, that NCBI never fully committed to PubMed Commons, and that this same attitude is apparent in their approach to archiving the data. I guess it was an interesting, if flawed experiment” (para. 13).

A return on innovation investment is typically determined by comparing the benefits of the new product or service to the research, development, and other direct expenditures generated in creating the innovation. The only evidence of such a cost-benefit analysis of PubMed Commons was the aforementioned declaration in the Feb. 2018 NCBI Insights blog post. My overall confusion about PubMed Commons and its sponsorship was heightened when

I recalled that Lipman initially stated that “this [PubMed Commons] is not an NIH project, except that NIH is allowing them to do it. We need to get the drive and direction from people who are going to be active in it” (Oransky, 2013, para. 23). In the end, the NIH considered PubMed Commons “a valuable experience in supporting discussion of published scientific literature” (NCBI Insights, 2018, para. 3).

Had project management been part of the plan for diffusion of the forum, evaluation would have included success criteria, a time-phased budget, scheduling, and milestone gates. Throughout this case study I wondered which sponsor(s) determined PubMed Commons’ value was solely based on participation. Recall that in 2016 Bastian pointed out that “PubMed Commons never needed to be huge, just open” (H. Bastian, personal communication, October 28. 2013). In her 2017 presentation, she reported that “the quality of comments is high; the use of the comments is low” (20:59). She had shared that a different evaluation method was needed.

> [PubMed Commons] needs grassroots thrust with research communities and slowly build the case for the value of commenting. Part of this is changing the perception of how you measure success. With regard to PubMed Commons, the number of comments isn’t the answer. (H. Bastian, personal communication, October 28, 2016)

In summary, staff investment appeared minimal. The forum obviously had initial drive and direction from invested scientists, but to what extent did it wane? Had PubMed Commons editors become distracted by the energy surrounding a growing debate over evolving peer review and publishing models? Was there a more needful project that required NCBI staff time and attention? I have been unable to find additional documentation detailing a cost-benefits analysis, project wrap-up, or lessons learned.

5.2.2.2 Technology

Rogers (1995) believed technology is often a driver of innovation, and the way it functions can influence adoption and continued use of a new idea, practice, or philosophy (p. 12). When examining factors that might have affected the forum's use, one cannot ignore issues related to the technological system. Recall the previous discussion about user experience and that technological functionality and accessibility should be developed in concert with the user perceptions about usefulness and how users interact with a system.

During the forum's development phase, Jonathon Dugan of PLoS labs was recruited to gather strategic advice from publishers and help design the system (Tibshirani, 2013, para. 7), which for all practical purposes worked. Throughout the closed pilot phase (June 17, 2013, to Oct. 22, 2013), "the [PubMed Commons] user group noted bugs and made a number of requests for modifications" (para. 1). This indicates that organizers recognized the need for a symbiotic exchange between Iterative- and Agile-level factors, as indicated on the forum innovation agility model. For example, user feedback from the social system in the Iterative layer caused changed in technology in the Agile layer, to improve the likelihood of adoption.

Authors of publications indexed in PubMed were encouraged to register after the forum moved to the open pilot phase. Initial dissatisfaction and frustration over the registration process caused NCBI "to investigate ways to open Commons up directly and automatically to more groups of published scientists" (para. 2). Embedding the PubMed logo, trending articles, and PubMed Commons "Reader Comments" search filter to the PubMed GUI and developing an altering system for informing authors when a comment was posted on their article's abstract (H. Bastian, personal communication, October 28, 2016). As a PubMed user, I thought the additions enhanced the forum's usability and its the presence, increasing the chances of diffusion by leveraging characteristics that could speed up the adoption process, like "visibility" and "communication" (Rogers, 1995, p. 17).

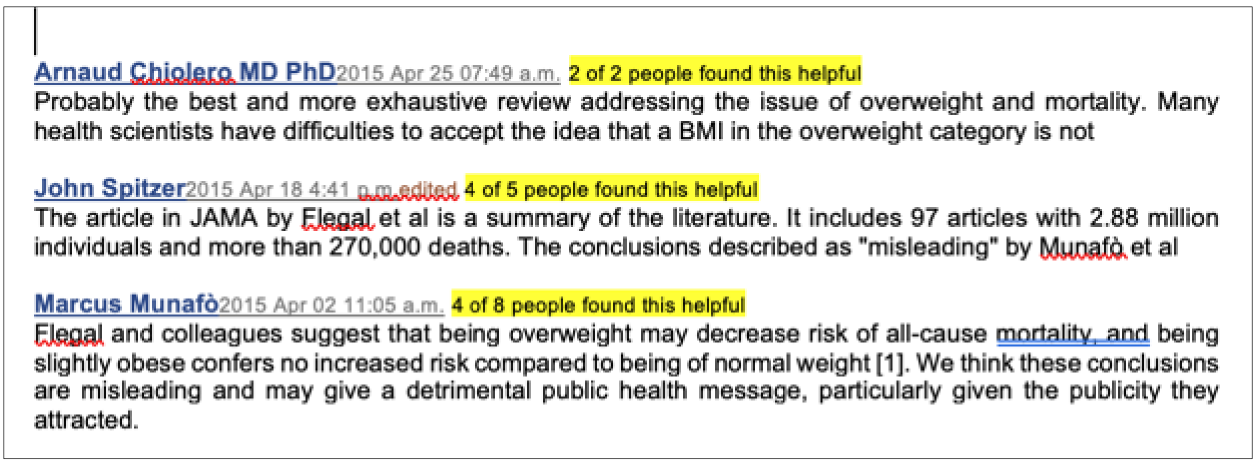
Prior to some of NCBI's technological updates, ambitious techies who were early adopters of the forum devised a number of workarounds to obtain forum information they were interested in. For example, before access to the "Readers Comments" search filter, PubMed users entered "has_user_comments[sb]" into the search box to retrieve articles with PubMed Commons comments (Lindebaum, 2014). Marie Ascher (2014) shared a method for setting up an author alert by using either the author's name (e.g., ascher mt [author] AND has_user_comments [filter]) or PubMed identification (PMID) (e.g., 23415612 [pmid] AND has_user_comments [filter]). Alf Eaton (2016) shared directions for writing a URL that returns JSON-formatted PubMed Commons data for a given PMID. Saunders (2016) expanded Eaton's work and wrote Ruby code for accessing all PMIDs with comments to fetch comment data and output a summary into a csv file. Such information sharing demonstrated that various social networks knew about PubMed Commons and had a vested interest as a user, reader, or steward of information.

An ongoing complaint about PubMed Commons was NCBI's standard API. Cockerill (2013) called it "very basic" and forecast that a "more extensive API [was] planned which would make it possible for publishers to incorporate PubMed Commons commenting more closely into their sites, perhaps entirely replacing publisher-specific commenting facilities" (para. 6). Saunders (2018) was not so optimistic the problem would be addressed, posting on various social media platforms that PubMed Commons needed “a real API.” After the forum’s discontinuation, he shared that this never happened.

High hopes for a rating system—helpful/not helpful—were dampened when PubMed Commons editors learned that people weren’t using it (H. Bastian, personal communication, October 28, 2016). Figure 5.4 shows comment helpful ratings. “4 of 8 people found this helpful” after Marcus Munafo’s comment means that four people found it helpful and four did not.

Figure 5.4

Healpful Ratings on PubMed Commons



Arnaud Chiolero MD PhD 2015 Apr 25 07:49 a.m. 2 of 2 people found this helpful
Probably the best and more exhaustive review addressing the issue of overweight and mortality. Many health scientists have difficulties to accept the idea that a BMI in the overweight category is not

John Spitzer 2015 Apr 18 4:41 p.m. edited 4 of 5 people found this helpful
The article in JAMA by Flegal et al is a summary of the literature. It includes 97 articles with 2.88 million individuals and more than 270,000 deaths. The conclusions described as "misleading" by Munafò et al

Marcus Munafò 2015 Apr 02 11:05 a.m. 4 of 8 people found this helpful
Flegal and colleagues suggest that being overweight may decrease risk of all-cause mortality, and being slightly obese confers no increased risk compared to being of normal weight [1]. We think these conclusions are misleading and may give a detrimental public health message, particularly given the publicity they attracted.

Recall in Chapter 4 that this feature, along with user IDs to disambiguate usernames; threaded and nested comments that showed which comments were replies to other comments; and information about comment moderation were not preserved in the final “commons_archive.csv” file (Saunders, 2018). This loss is indicative of disappointment over promised perpetuity of forum data. If someone did not preserve the forum data like I detailed in Chapter 4, future investigations about critical factors that affected PubMed Commons are not possible.

Fortunately, organizations like Europe PMC (2018) and hypothes.is (2018) have preserved PubMed Commons comments. While PubPeer has integrated most comments into its interface, not all are available.

One shortcoming forum users expressed was that the PubMed search engine does not provide access to full text. This is important because PubMed users without paid access to publications would have to evaluate an article receiving comments based solely on the article abstract. Furthermore, scientists who did have access might be limited by time and read only the abstract and comment instead of the entire article.

> There is ample evidence to show that what gets into the published biomedical literature is generally poorly reported in a number of important ways (search Google for "pubmed papers on reporting quality"), so it makes far more sense to focus on quality of what gets published, rather than commenting on abstracts of poorly reported publications. (G. Stevan Bova as cited in Oransky, 2013, comment 10)

Ironically, PubMed Commons' technology did not keep pace with advancements in the shift to Internet-based publishing. Traditional bibliometrics (e.g., citations and journal impact factor) that measure research influence and impact continue to be supplemented, perhaps eventually supplanted, by attention scores reported by companies like Altmetric and Plum Analytics. Altmetrics "donut" scores include measures from post-publication peer review forums, of which PubMed Commons would be considered (Digital Science, n.d.). This attention would have provided ongoing promotion to the forum, thus increasing the likelihood of an increase in rate of adoption.

Overall, the technology that supported PubMed Commons lagged other systems.

My environmental scan of commenting forums like PubMed Commons showed that both BMC and PubPeer, among others, had developed technologies to highlight the existence of comments on their websites—even on PubMed (see Figure 5.5). PubPeer engineered a

Google Chrome plug-in for its system, as shown in Figure 5.6 (Chrome Web Store, n.d.). On his *Retraction Watch* blog, Oransky (2017) informed his readers that

> Richard Smith and Steph Smith-Unna recently created a great API to easily access publication events, and we [PubPeer] are using it to more systematically pull information into PubPeer, make it searchable, and insert it into the appropriate timepoints of the PubPeer commenting timelines. (Brandon Stell, para. 10)

Such innovations sparked communication among the growing social system concerned with after-publication commenting and increased visibility to alternative forums. It is possible innovative system add-ons could have raised the user experience value of certain websites and become part of the way potential users assessed one system over another or tried out and observed a system. All of these elements could affect forum adoption rates (Rogers, 1995, pp. 16-17).

Figure 5.5

API-generated Notice for PubMed Commons on BMC GUI

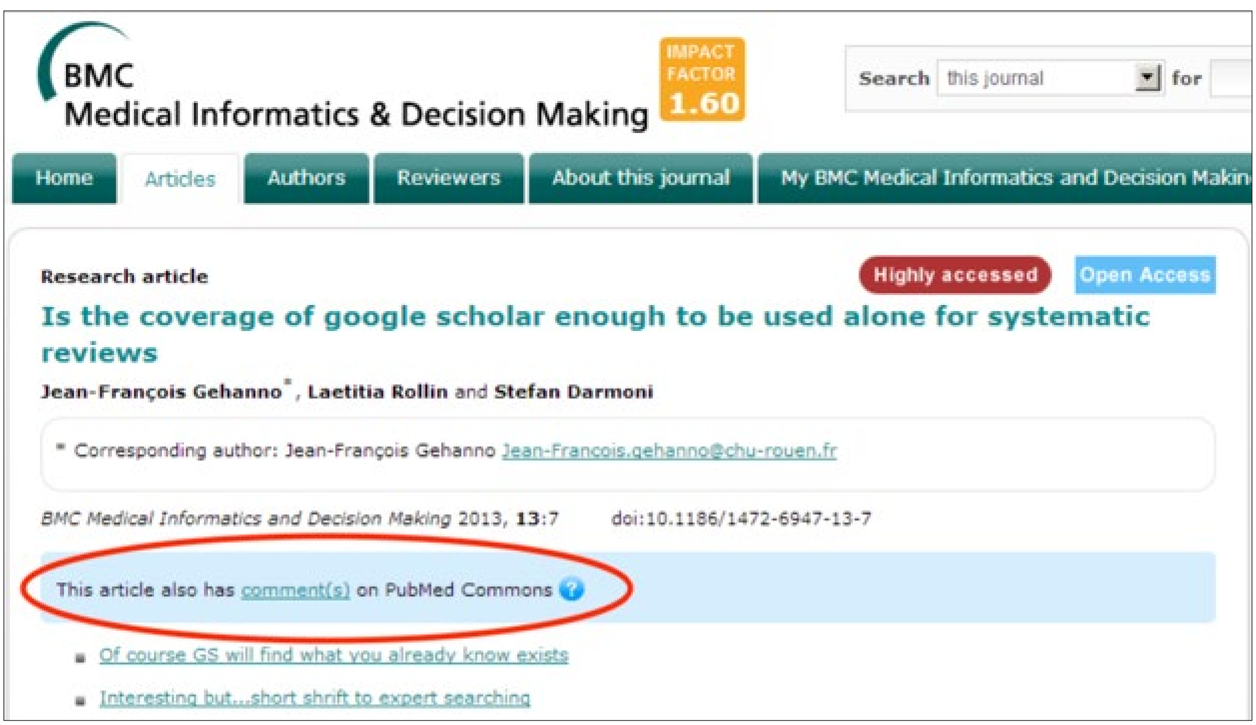
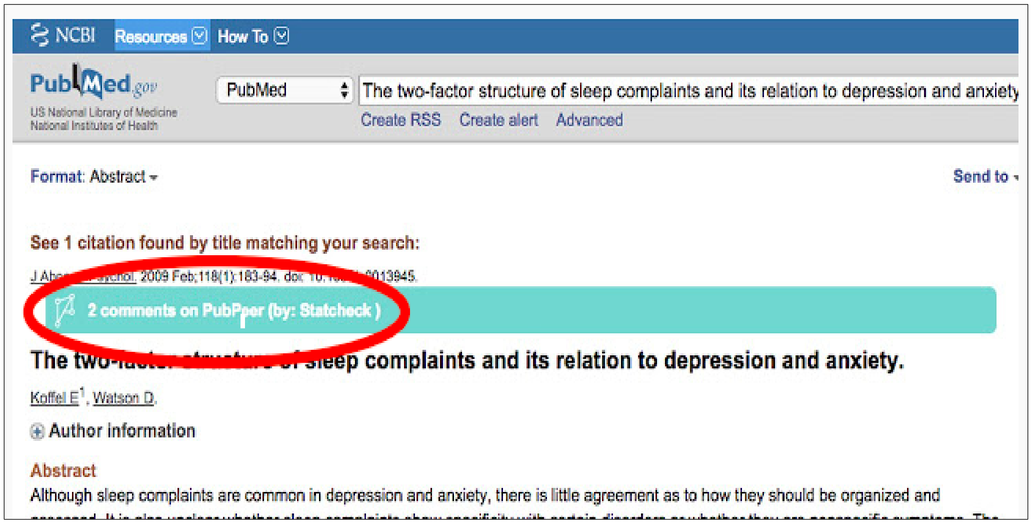


Figure 5.6

API-generated Notice for PubPeer on PubMed GUI



Delays in technological innovations have been reported as symptomatic of government-sponsored websites. At a Harvard Kennedy School forum Harvard Law School Professor Susan Crawford explained that large tech companies (e.g., Amazon, Berkshire Hathaway, JP Morgan Chase) have their own infrastructures and are developing things like private internet services and health care systems.

> At a time when the U.S. subway system is falling apart, Amazon is building a heliport. Health care, transit, communication … these are all essential for America, yet these giant companies can build around them. It illustrates something profoundly wrong with American government. (Milano, 2019, para. 3, 4)

The need for working and updated technology addresses many theoretical elements of DOI that could affect diffusion speed and rate of adoption (e.g., relative advantage, trialability, observability, communication, and time) (Rogers, 1995, pp. 16-17).

5.2.2.3 Championing

Innovations are more quickly adopted when they are championed by individuals who share core concepts of a project with colleagues or other members of a social system. "[Champions] make or break long-term success" (Wells, 2020, para. 6). They support the innovation, identify novel ways to bring value to users, and can be the voice of reason for innovators who don't want to listen to people outside their organization (para. 7). I maintain that Bastian was a visible PubMed Commons champion.

Recall that NCBI Insights (Nov. 1, 2013) told its readers "PubMed Commons is an un-moderated commenting system." Although Bastian, Vaught, and Lipman served as moderators, they were never publicly recognized in this role. Lipman decided when comments were removed if a commenter violated objective guidelines (H. Bastian, personal communication, October 28, 2016). Findings in Chapter 4 showed how Bastian was a central commenter on PubMed Commons, posting more than 71 comments on 60 articles and interacting with 51 commenters. Her posts were characterized as discussions. Based on social network analysis data, she was the most connected person on PubMed Commons. In this chapter, I've shown how Bastian was an avid supporter of the forum, promoter of after-publication commenting, and voice for PubMed Commons. She visibly filled the role of a champion by modeling the way she was using the innovative forum for open discussion about publications so that others could imitate her, as Tarde advocated (Kinnunen, 1996). According to Rogers (1995), this observability would aid PubMed Commons in being adopted more rapidly (p. 16). By reading Bastian's comments, potential adopters could also consider the forum's relative advantage over alternatives, as well as evaluate the compatibility of comments and discussions with their existing values, past experiences, and needs related to after-publication commenting (pp. 15, 16). The opportunity and, frankly, responsibility for others to serve as champions (e.g., PubMed

Commons innovators, organizers, and Team Leads) was noticeably missed.

Had the forum been iteratively evaluated, agile modifications to how organizers viewed and utilized the role of moderator could have positively acted not only forum adoption, but also sustained use.

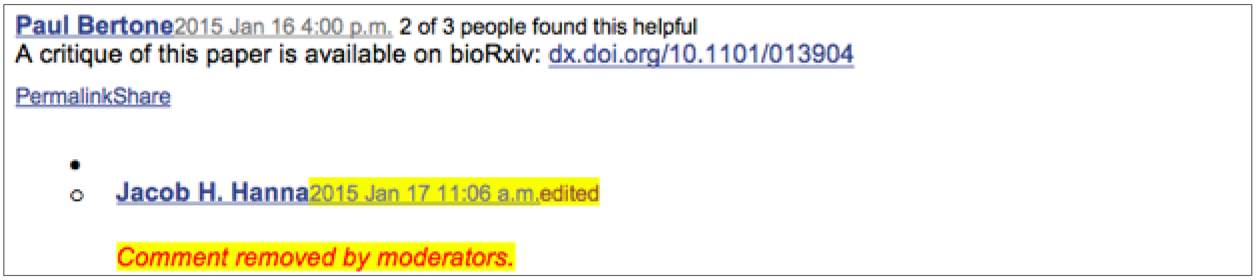
> Special status is conferred upon those who actually regulate the communication activity [of an online forum]. Moderators/administrators are vested with special authority to change the structure of signs by editing, moving or deleting messages (and thereby acting as 'gatekeepers')… as well as by imposing disciplinary penalties for inappriorate behavior, ranging from warnings to complete account deletion. (Bylieva, Lobatyuk, & Safonova, 2019, p. 334)

Researchers have identified tactics for moderators to increase online forum adoption and enhance the development of a social system among users. In addition to modeling appropriate behavior and netiquette guidelines, moderators can set a tone for discussions, post messages to enforce rules of conduct, assure a safe commenting environment, deal appropriately with unacceptable posts, douse flaming, answer questions, share their expertise, and moderate discussions (Fronzetti Colladon & Vagaggini, 2017; Heckman & Annabi, 2006; Mokoena, 2013). Early forum adopters often interact with the moderator for the purpose of obtaining more information (Fronzetti Colladon & Vagaggini, 2017, p. 1295), which would move an early adopter from the first stage to the second stage of the adoption process (Kaminski, 2011, p. 4).

Instead of encouragement from moderator-champions, PubMed Commons users could visibly see that the forum did indeed have moderators, contrary to the *NCBI Insight* post mentioned above. Adopters and others observing comments on abstracts appearing on the PubMed GUI, repeatedly saw "Comment removed by moderators" which potentially added confusion to how the forum was managed and threatened trust (see Figure 5.7).

Figure 5.7

Comment Removed by Moderators on PubMed Commons



Results from my quantitative content analysis reported that 93 comments had been removed by moderators. (See Chapter 4 for detailed information about moderating on PubMed Commons). Early adopters reported in social media posts they had been advised to edit a post when it violated a guideline, or the post would be removed. Some adopters shared they were completely banned from the forum. These mounting negative testimonials might have discouraged forum adoption or affected potential adopters' evaluation of whether using PubMed Commons was advantageous over other alternatives for commenting about published articles (Rogers, 1995, p. 16).

About moderating, Bastian admitted that "some people might disagree with the threshold of moderating" (H. Bastian, personal communication, October 28, 2016).

> We think we have a sustainable infrastructure that minimizes potentially harmful use of the forum—that was one of the main criteria that the NIH had for this. Some will criticize the way we've achieved this, but essentially, we think we have and the feedback seems to be by and large that people think that. (Bastian, 2016, 20:50)

In contrast to too much interference on the forum, there were noticeable oversights in removing posts that could be considered spam (e.g., Kath Wright's promoting InterTASC Information Specialists' Sub-Group). Again, adopters' perceptions and forum diffusion could have been affected. Effectively curating the forum would signal that the forum had champions who were guardians of the forum, not only monitoring for possible harmful posts, but also reducing the noise of irrelevant messages.

The decision of PubMed Commons organizers to downplay the visible role of an online moderator-champion on the forum might have been profound. Fronzetti Colladon & Vagaggini (2017) found that in a forum's early stages, moderators play an even more important role (p. 1294), and that removing a moderator has a significant impact on network connectivity and the shared content (p. 1287). Recall that PubMed Commons' network cohesion was alarming low to the extent that the UCINET cohesion report produced a warning message that the network was disconnected. Bastian, however, was highly connected, brokering connections and serving as a hub that created pathways between commenters. My data showed that while the average, normalized betweenness (i.e., the number of times a commenter acts as a bridge along the shortest path between two other commenters) was 0.029, Bastian's was 3.285, some 113 times the network's mean.

It could be that having moderator-champions take a more active role by utilizing evidence-based speech acts, PubMed Commons rate of diffusion could have sped up, and its adoption broadened. I recommend an investigation of this theory as an area for future research.

5.2.3 Iterative Layer

Based on my case study of PubMed Commons, the four main elements Rogers (1995) identified as drivers for diffusion of innovation—innovation itself, communication, time, and social system—could have been used as a guide for evaluating the PubMed Commons forum in terms of rate of adoption, forum diffusion among eligible researchers, and user perceptions. I have included the four in the Iterative Layer of the forum innovation agility model, albeit the terminology was modified to better fit with diffusion of an online forum. These changes are explained below.

Optimally, each factor should be part of an overall project management strategy, and after initial forum implementation, they should be proactively and repeatedly (i.e., iteratively) measured and analyzed as part of milestone gates, quality checks, risk management/mitigation, and timeline scheduling. Doing so would inform project innovators and organizers about forces happening outside their control (i.e., "that's outside my boat) so they are able to make informed, agile adjustments to factors inside their locus of control. It is important for organizers to keep in mind that the objective is successful adoption and diffusion of the innovation among targeted stakeholders (i.e., winning the medal).

5.2.3.1 Forum Itself

Rogers (1995) believed technology is often a driver of innovation, and the way it functions can influence adoption and continued use of a new idea, practice, or philosophy (p. 12). In the case of PubMed Commons, the online forum embedded in the PubMed GUI is the innovation under study. Although online forums have been in existence for over 40 years (Driscoll, 2016), "the perceived newness of the idea" (Rogers, 1995, p. 11) is what makes it an innovation. As explained above, there had never been an online forum for open discussion about published scientific literature on such a grand scale as PubMed Commons.

Research questions to evaluate the innovation (i.e., forum) might investigate how early adopters differ from late adopters, how perceived attributes (i.e., relative advantage or compatibility) affect the rate of adoption, and what factors increase the rate of adoption

(Rogers, 1995, pp. 11, 12). Chapter 4 of this dissertation details four methods I used to evaluate PubMed Commons: 1) quantitative content analysis, 2) user demographics, 3) qualitative content analysis, and 4) social network analysis. The PubMed Commons team collected data during a three-month period in 2014 and found that there were more single than threaded comments; 70% of publication with comments centered on primary research—over reviews, commentaries, and methods of resources; most comments link to literature or other resources; and most comments are posted to recent publications (see Appendix G). Vaught, Jordan, and Bastian (2016) collected usage data for comments from Jan. 2016 to Dec. 2016, discovering that most individual members of PubMed Commons had not commented and that only a small number of members accounted for a considerable portion of the comments. They also found that a) comments rarely included discloser about conflicts of interest; b) geographical distribution of commenters was not representative of authors in biomedical literature; and c) women were underrepresented (para. 4). In 2017, Vaught, Jordan, and Bastian published/presented about data they collected regarding editorial expressions of concern in PubMed and PubMed Central. This research focused on post-publication activity that was not specifically related to the evaluation of the PubMed Commons forum.

Other tangible measures of forum adoption included number of registered users and posted comments. This information was communicated by various PubMed Commons organizers in different communication channels (see Table 5.2).

Without question, an online forum is a complex organism of study. There is “almost an unlimited amount of material for analysis” (Holts, Kronberger, & Wagner, 2012, p. 56). Suggestions for evaluating an online forum include user surveys and interviews, focus groups, usability testing, netnography, and web analytics, among others. Because online forums are an element of social media technologies (i.e., "online media that stimulates participation, openness' conversation, connecters and sense of community" [Saravanakumar & SuganthaLakshmi, 2012, p. 4444]), SNA is a fitting method for discovering helpful information that could trigger mitigating factors to achieve projected adoption and use.

Table 5.2

Rate of Adoption Reported by PubMed Commons Team

Date	Members	Journal Clubs	Comments	Publications
Dec. 17, 2015	9,500	20	4,000	3,300
Oct. 31, 2016	10,632	24	5,739	4,595
Nov. 21, 2016	—	—	80% single; 12% threaded	1,400 in 2016
Dec. 31, 2016	10,736	24	5,483 (2014 to 2016)	—
Mar. 31, 2017	10,962 w/ 1,637 commenters		6,372 on 5,078 publications	—

Observable levels of commenting frequency, directionality, and intensity (i.e., who comments on whose post and how often) can be examined to determine connections among users and which individuals hold key positions on the forum (i.e., hub, broker, gatekeeper, isolate). SNA research focuses on relationships, similarities in behavior, social relations, and flows (Haythornwaite, 1996). Repeated SNA evaluations of the same forum can provide information about rate of adoption and the diffusion of the forum in a defined community. Comparing an SNA with another of similar size would improve interpretations of density, component, and cohesion measures. One exciting development in SNA research is the ability to leverage newly developed analytical tools to predict the diffusion of an online forum, which, again, could inform organizers to implement agile factors and mitigate diffusion that falls short of threshold expectations (Barabasi, 2012; Gabbriellini, 2014; Yan, Tsekenis, Barzel, Slotine, Liu & Barabasi, 2015). Regardless of method, evaluating the forum (i.e., innovation itself) is essential for ensuring project management strategies are met and that the diffusion is progressing as projected.

5.2.3.2 Promotion

Rogers (1995) identified communication as one of the four main elements that "are identifiable in every diffusion research study, and in every diffusion campaign or program" (p. 10). He conceived communication as "the process by which participants create and share information with one another in order to reach mutual understanding" (p. 17). As described above, organic communication about PubMed Commons (i.e., interpersonal in social media, NCBI/PubMed Commons blog posts and YouTube videos) provided mixed results for demonstrating relative advantage and compatibility, reducing complexity, or offering a better understanding of the forum. Forum rate and speed of adoption might have benefitted if additional, strategic communications focused on mass media channels to create "awareness-knowledge" about the forum (p. 18). Likewise, understanding the value of a higher level or intensity of communication—to the point of "selling," might have prompted PubMed Commons organizers to recognize the need for regularly scheduled forum promotions.

Table 5.3 provides a calendar of communications about PubMed Commons championed by its editors, moderators, innovators, and sponsor organization leaders. Asterisks indicate when the terms post-publication peer review were used. As mentioned earlier, only seven blog posts were made on NCBI Insights, and only eight blog posts were made on the PubMed Commons Blog.

Table 5.3

Communication Calendar by Innovators and Organizers

Date	Resource	Title
Oct. 22, 2013	NCBI Insights blog	PubMed Commons: A new forum for scientific discourse
Oct. 22, 2013	Bastian, H. *Scientific American* blog	Science Buzz and Criticism get a Powerful Boost*
Oct. 22, 2013	Eisen, M.	PubMed Commons: Post publication peer review goes mainstream*
Oct. 22, 2013	Tibshirani, R.	PubMed Commons: A system for commenting on articles in PubMed
Oct. 23, 2013	NCBI Insights blog	Joining PubMed Commons: A step-by-step guide
Nov. 1, 2013	NCBI Insights blog	Early developments in the PubMed Commons pilot
Nov. 14, 2013	PubMed Commons blog	Welcome!
Nov. 26, 2013	PubMed Commons blog	PubMed Commons going public soon
Dec. 13, 2013	PubMed Commons blog	Comment search and alert: A PubMed Commons guide
Jan. 30, 2014	NIH Public Access	NIH plans to enhance reproducibility
Apr. 22, 2014	Goodman, S. YouTube.	Prof. Steven Goodman on PubMed Commons*
Aug. 5, 2014	Collins, F. NIH Director's Blog	PubMed Commons: Catalyzing scientist-to-scientist interactions
Nov. 14, 2014	*Nucleic Acids Research*	Database resources of the National Center for Biotechnology Information by NCBI Resource Coordinators
Nov. 2014	Bastian, H. *PLoS Medicine*	Stronger Post-Publication Culture is Needed for Better Science
Dec. 17, 2014	PubMed Commons blog	Introducing PubMed Commons Journal Clubs
Feb. 17, 2015	NCBI Insights	Professors: NCBI can help you streamline your teaching and research efforts
May 6, 2015	PubMed Commons blog	Signposts from research to resources

(table continues)

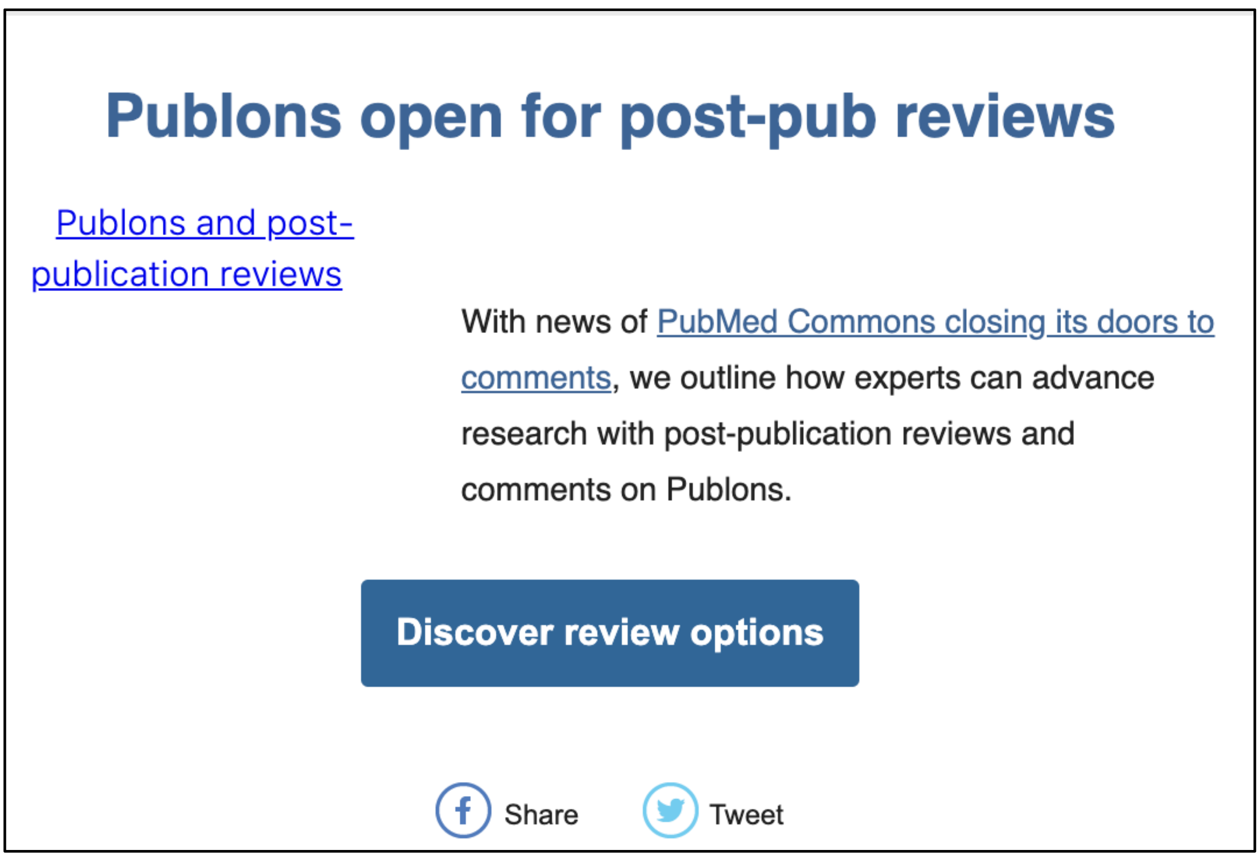
Date	Resource	Title
Dec. 17, 2015	PubMed Commons blog	Commenting on PubMed Commons: A successful pilot
Dec. 17, 2015	NCBI You Tube Channel NCBI Minute webinar	Quick introductions to NCBI resources
Feb. 11, 2016	AAAS Meeting	Post-publication activity on PubMed Commons
Apr. 21, 2016	NNLM YouTube channel	PubMed Commons has unveiled a new look
Oct. 22, 2013	NCBI Insights blog	PubMed Commons: A new forum for scientific discourse
Nov. 21, 2016	PubMed Commons blog	PubMed comments and their continuing conversations
Mar. 15, 2017	Bastian, H. NCBI webinar	Evaluation of Post-Publication Activities in PubMed
Mar. 21, 2017	Bastian, H. NCBI YouTube Channel	Post-publication peer-review and certificate systems. Disclosure that she is PubMed Commons lead editor
Apr. 6, 2017	PubMed Commons blog	Authors altering readers via PubMed Commons
May 28, 2017	PubMed Commons blog	Critiquing systematic review search strategies on PubMed
Jul. 5, 2017	PubMed Commons blog	Collaborating to bring journal clubs to PubMed Commons: A librarian's perspective
Jun. 30, 2017	NCBI Insights blog	July 12th NCBI Minute "Crowdsourcing post-publication comments: How you and your journal club can contribute using PubMed Commons"*
Jul. 26, 2017	Bastian, H. NCBI YouTube Channel	NNLM Resource Picks. Understanding systematic reviews and more at PubMed Health
Jul. 12, 2017	Vaught, M. NCBI You Tube Channel	PubMed Commons for journal clubs and authors NCBI Minute webinar
Jul. 17, 2017	NCBI Insights blog	New video on the NCBI YouTube channel: How you and your journal club can contribute using PubMed Commons
Feb. 1, 2018	NCBI Insights blog	PubMed Commons to be discontinued

* Uses terms "post-publication, peer, review"

Seeing and treating communication as a marketing strategy, would have brought into play purposeful consideration for competitive intelligence, differentiation, scheduled messaging, and measuring the effect of impressions on project efforts to diffuse forum commenting among PubMed Commons authors for the purpose of open discussions about scientific publications. Competitive intelligence would have identified that alternatives to PubMed Commons were available for commenting about publications (e.g., PubPeer, *Retraction Watch*, F1000, arXiv and bioRxiv, Cochrane Journal Club, journal commenting systems like BMJ Opinion, Bepress (purchased by Elsevier, later becoming Publons), TrueReview, OpenReview to name a few). The systematic collection and analysis of data about these other forums would reveal ways to differentiate PubMed Commons from alternatives. Marketing strategies could develop campaigns for communicating relative advantages of the forum over others and articulating ways in which commenting was compatible with users' existing scholarly communication behaviors.

Figure 5.8

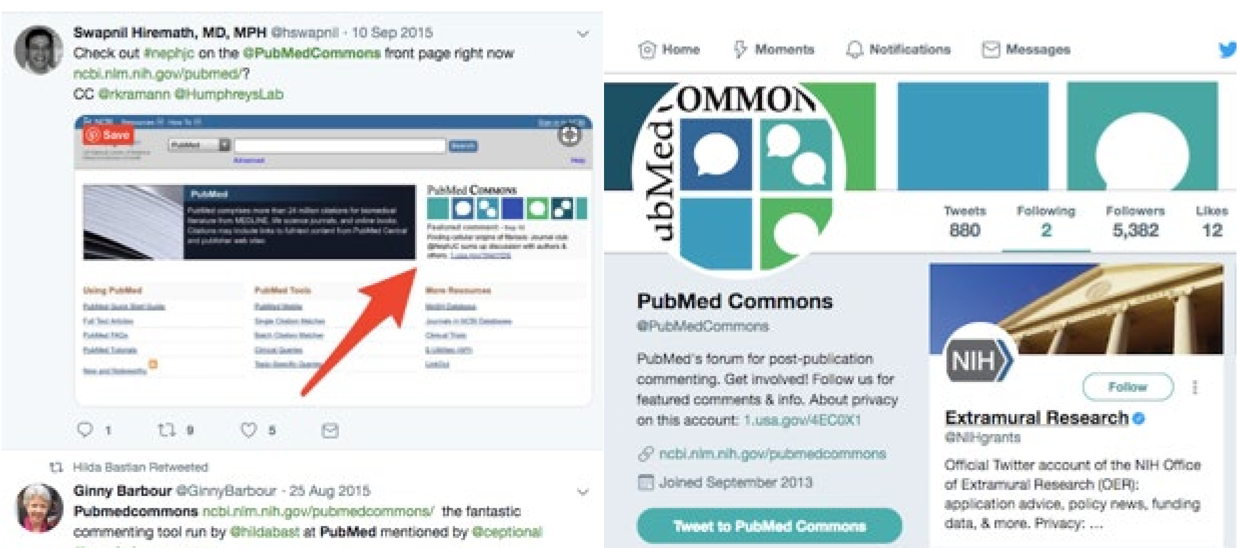
Elsevier Email Solicitation to Former PubMed Commons Users



Bastian speculated that commenting on PubMed Commons might be intimidating because "it's PubMed," meaning that commenters believed their comments should match the level of scholarly writing indexed in PubMed (H. Bastian, personal communication, October 28, 2016). Accordingly, the forum's organizers would have to weigh this impression of the forum with a more commercialized approach to compete for users' participation and the time they would invest in writing a comment. By contrast, organizations like Elsevier, which calls itself "the modern publishing business," can operate under a different set of promotion guidelines. As a global business with a 150 history, it can be more aggressive in its marketing. For example, in late February 2018 after the discontinuation of PubMed, Elsevier executed an email campaign and strategic communications to promote its Publons service as a replacement for PubMed Commons (see Figure 5.8).

Figure 5.9

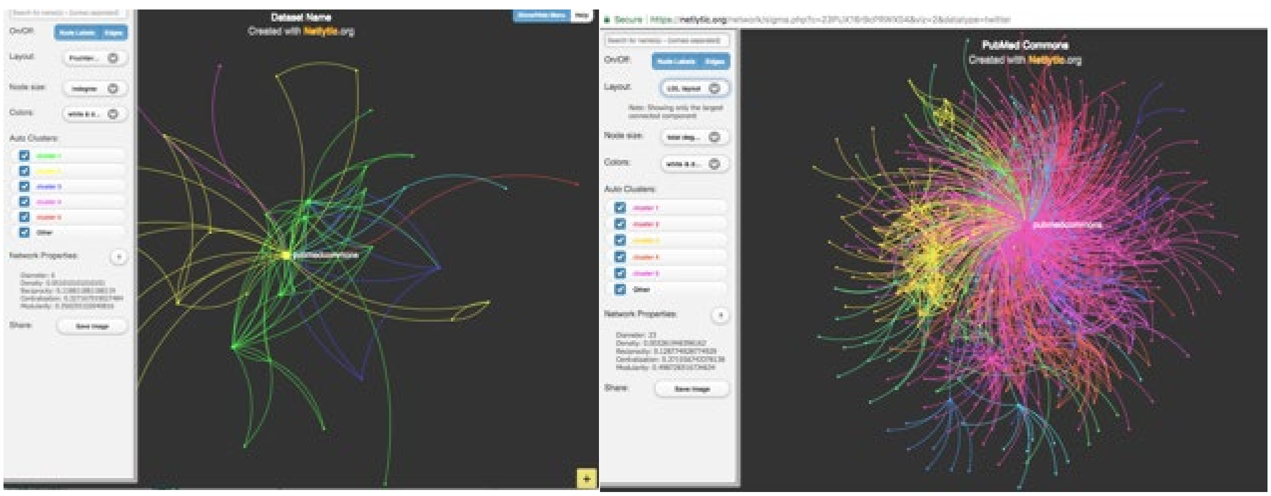
PubMed Commons Presence on Twitter



Abundant literature reports best practices for marketing, promotions, and strategic communications. One example is the "CDC Social Media Tools, Guidelines, & Best Practices" webpage, which links policies, guides, security mitigations, and a toolkit to assist the CDC and other federal, state, local agencies, and private organizations (CDC, 2021, para. 1). The PubMed Commons team effectively used Twitter to regularly communicate with forum users, adopters, and potential adopters (see Figure 5.9). On Sep. 10, 2105, PubMed Commons announced new PubMed GUI features on their Twitter account. At the time the PubMed Commons Twitter account was shut down, there were 5,382 followers and PubMed Commons editors made 880 tweets.

Figure 5.10

*Network Sociograms Comparing PubMed Commons Tweet*s



A Netlytics analyses of the PubMed Commons Twitter account performed in Nov. 2015 (see Figure 5.10, left) and again in Jan. 2018 (see Figure 5.10, right) showed a growth in the network size of tweeted and retweeted tweets about PubMed Commons. On Nov. 15, 2015, the PubMed Commons Twitter "Name Network" (i.e., who mentions whom) had 32 nodes with 101 ties. Its "Chain Network" (i.e., who replies to whom) had 9 nodes and 23 ties. On Jan 11, 2018, the PubMed Commons Twitter "Name Network" had 552 nodes with 1,766 ties. It's

“Chain Network” had 161 nodes with 674 ties. Overall interaction increased 17-fold, apart from the number of relationships in the Chain Network, which grew 29-fold.

When comparing sentiment word clouds and the counts for popular terms tweeted on Nov. 11, 2015, and Jan. 1, 2018, also highlights increased interaction. The subject matter of comments remained focused on scholarly communication (e.g., review, study, systematic, journals, peer, discussion, letters, paper, and editor). PubPeer, an alternative post-publication commenting forum, appeared in both word clouds (see Figures 5.11 and 5.12). The PubMed Commons teams could have exploited similar web analytics technologies to monitor their social media promotions and quicky discover what people were discussing.

Figure 5.11

Sentiment of Tweets on November 15, 2015

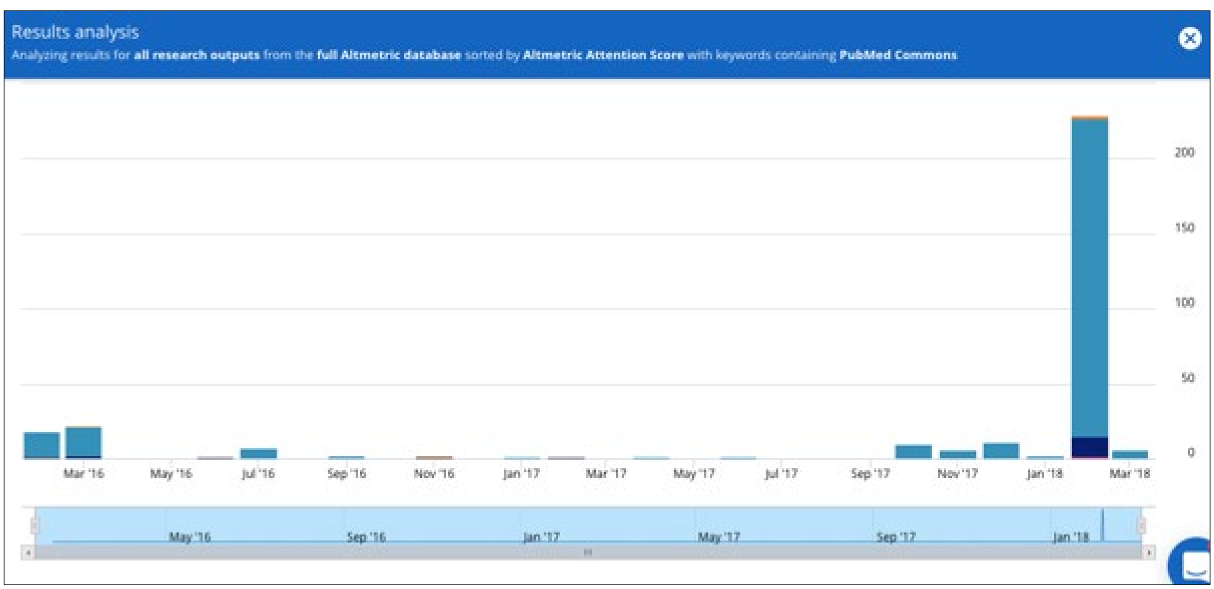
Figure 5.12

Sentiment of Tweets on January 11, 2018

Altmetric Explorer provides insights from the Altemtric database about attention data on scholarly content. The Altmetric Attention Score is calculated from all research outputs, including relevant online discussions, policy documents, mainstream news outlets, post-publication peer-review forums, social media and blogs, Wikipedia, and online reference managers. Most attention for PubMed Commons came from Twitter. The spike in Feb. 2018 reflects the large Twitter attention PubMed Commons received when its discontinuation was announced (see Figure 5.12). The turquoise color represents attention from Twitter. The dark blue is Facebook attention. The gold is from blogs, the dark red from news, and plum from Google+. The figure clearly shows that PubMed Commons did not receive much attention between Feb. 2016 and Jan. 2018.

Figure 5.13

Attention Scores for Altmetric Explorer Search for "PubMed Commons"



5.2.3.3 Time

"The inclusion of time as a variable in diffusion research is one of its strengths, but the measurement of the time dimension (often by means of the respondents' recall) can be criticized" (Rogers, 1995, p. 20). Time measures can report a) the span between an individual's awareness-knowledge of the innovation and his or her adoption of it or b) the difference among individuals who adopt the innovation (e.g., innovators [2.5%], early adopters [13.5%], early majority [34%], late majority [34%], laggards [16%]) (p. 262).

Another way to measure the time dimension is the rate of adoption (i.e., the number of members of the system that adopt the innovation in a given period of time). For PubMed Commons, this last measure makes sense for organizers to consider and, as previously mentioned, mitigate lower than expected adoption rates by changing agile factors. For example, surveys could be sent to targeted users and adopters, asking about their user experience. Would changes in guidelines or interface design or accessibility cause more people to adopt the innovation or adopters to use the innovation more often? Changes in technology could include added features or redesign. Perhaps a new sponsor group or influx of financial support would provide needed salary for an added employee, boost in promotions, travel funds to conference, or monies to pay open access article processing charges (APCs). Given that a forum is moderated, champions could be encouraged to intensify their efforts to forge online relationships or prompt discussions.

NIH directors began their sponsorship of PubMed Commons believing that funded research needs transparency, and the forum did not need to have a large number of adopters. The forum just needed to be open (H. Bastian, personal communication, October 28, 2016). The NIH discontinued PubMed Commons saying,

> The service was first introduced as a pilot project in the fall of 2013 and was reviewed in 2015. Despite low levels of use at that time, NIH decided to extend the effort for

> another year or two in hopes that participation would increase. Unfortunately, usage has remained minimal, with comments submitted on only 6,000 of the 28 million articles indexed in PubMed. While many worthwhile comments were made through the service during its 4 years of operations, NIH has decided that the low level of participation does not warrant continued investment in the project, particularly given the availability of other commenting venues. (NCBI Insights, 2018, para. 3, 4)

According to this announcement, the time dimension measured was 24 months. The expected rate of adoption was not indicated. Although PubMed Commons editors believed the common purpose of the forum was creating an online space for "open discourse about published articles" (Collins & Tabak, 2014; H. Bastian, personal communication, October 28, 2016), the forum's discontinuation implies there was an expected rate of adoption within a specific amount of time. This message was not publicly communicated. Had the rate been promoted publicly, strategies could have been implemented to encourage adoption (e.g., a thermometer graphic on the PubMed Commons blog that showed adoption rate and motivated individuals to adopt the forum; a counter on the PubMed GUI interface with adoption numbers in real time).

5.2.3.4 Social System

Chapter 4 and my discussion in the sponsorship section above demonstrated that PubMed Commons had impressive sponsorship that could very well translate into a social system supporting the diffusion of the forum. PubMed Commons organizers engaged in communication, although it was infrequent, characterized as inconsistent messaging, and not to the intensity of promotion. There is no evidence of the inclusion of a time dimension as part of iterative processes to evaluate the forum.

NIH sponsorship of PubMed Commons afforded innovators and organizers the ability to position forum comments on the PubMed GUI, granting it visibility to achieve a healthy social presence among potential users throughout the world. Unfortunately, this did not facilitate the formation of a much-needed social system to encourage adoption of the forum. Rogers' (1995) conceived a social system as "a set of interrelated units that are engaged in joint problem-solving to accomplish a common goal" p. 23).

NCBI sponsors trusted that connected (i.e., interrelated) forum innovators, Team Leads, and the 250 scientists recruited as early adopters would coalesce into a social system that would power the diffusion of the highly anticipated post-publication commenting forum (i.e., solve the problem of having a central place to discuss scientific publications).

Artifacts and cited communications throughout this chapter have demonstrated that PubMed Commons innovators, organizers, and early adopters were part of a variety of strong offline networks that one might expect could transition into online connections. During the open beta pilot phase, there was an observable group of top-level, widely respected, highly visible PubMed Commons supporters who were networked among government agencies, health and patient organizations, research centers, the publishing industry, journalism and social media outlets, and foundations across the globe. Furthermore, the fledgling forum had the blessing of the NIH, whose $41.7 billion annual investment in medical research directly impacts more than 300,000 researchers at 2,500 American universities, medical schools, and other research institutions in every American state (NIH, 2020, para. 1, 2). The reach of this scientific community is massive. One must remember, however, that to comment on PubMed Commons, an individual had to be an author of an article indexed in PubMed. This guideline limited the ability of individual to network into an online social system. Results from the social network analysis reported in Chapter 4 proved that a social system did not form. On the contrary, the forum was a network of independents with individual agendas on a spectrum of convictions about post-publication commenting.

The social structure of a system affects the DOI in many ways. “The social system constitutes a boundary within which an innovation diffuses” (Rogers, 1995, p. 24). Perhaps the social system of scientists working around the globe with articles indexed in PubMed was too large for diffusion. Would adoption have increased if the forum allowed adopters to self-select into groups for social support or tag comments to narrow the overall scope of “comments on PubMed Commons?” Such agile technical features might have made a difference in user experience or helped overcome concerns about what constituted as a “peer.”

The effects of norms within a Community of Practice, the roles opinion leaders play, project management decisions, and intended and unintended consequences of the innovation are considerations in the diffusion process (p. 24). Norms entail behaviors and structure implies predictability (p. 24). It could be that the forum was too open and the hierarchical structure too flat for a group of scientists accustomed to working in a bureaucratic organization and publishing their work in a system built on traditional peer review. The concept of posting comments after publication might have been too far outside habitual patterned social relationships.

Communication norms are also part of a social structure. Considering that issue of evolving perceptions about peer review distracted the existing social system and individuals willing to join the system from the task at hand—diffusing the PubMed Commons forum—the innovation could have been too drastic and overwhelming. Recall from Chapter 2 that the NLM introduced a biological preprint service via Information Exchange Groups in the early 1960s, which attracted over 3,600 participants and produced over 2,500 documents before it was discontinued in 1967. The fact that the NIH referred to PubMed Commons as a "valuable experiment" (NCBI, 2016, para. 1) implies that the NLM has a history of innovating technologies that are simply before their time. Preprint forums are growing in popularity and offer researchers a process to share scientific manuscripts in an online public repository before peer review. This Internet-hosted meeting of the "Invisible College" promotes a collegial way offer advice before publishing, not after, like PubMed Commons. Launching 25 years after the NLM's preprint service was discontinued, arXiv is celebrating 30 years of open science this year, boasting 1,939,272 scholarly articles in a range of scientific fields (Cornell University, n.d.). BioarXiv launched in 2013 and medRxiv in 2019. PrePubMed, which is not affiliated with NCBI or PubMed, indexes preprints, aggregating access for users and eliminating the need to visit the growing number of sites devoted to this scholarly communication practice.

A shift in the publishing paradigm is afoot. NLM's PubMed Commons experiment might have filled the need for a visible change agent, maybe just too soon.

5.2.4 Outside Layer: Adoption in Action

The previous section is a fitting segue to the Outside layer of the forum innovation agility model, primarily because the forthcoming change in scientific publishing was the single distraction that continually tipped the boat for people in charge of steering the direction of PubMed Commons. The idea of a revolution in publishing was already out there, and because the forum was envisioned as a tool in the process, its purpose became muddled and intertwined with the larger scope of changes in scientific publishing.

Project managers refer to such a loss of focus as scope creep. “Scope creep is a dreaded thing that can happen on any project, wasting money, decreasing satisfaction, and causing the expected project value to not be met” (Larson & Larson, 2009, para. 1). It is not unusual for scope creep. Working on unintended features of a project, in this case spending time composing comments, writing blog posts, collecting data, researching related topics, writing articles, making presentations—whew—individuals devote their time on things outside their boat, in this case the larger issue of changes in scientific publishing. Typically, the clock (rather stopwatch) doesn’t stop on the project’s original purpose and timeframe. This means that certain aspects don’t get completed. Being aware of strategies to combat it are helpful. First, team members must remain focused on the common purpose. If necessary, separate groups should write down the purpose and how their role fits their contributions. Second, purpose statements should include features that are inside their boat and outside their boat. Project managers do this with the WBS. Third, the project should have clear, complete, and concise requirements. Take time to plan them, name them, and focus on them. Fourth, there should be a completion date. This should be tied to milestone gates so that mitigating changes can be made to get the project back on course. Fifth, follow a model or diagram as a visual aid to clarify the project flow and facilitate effective sharing of perceptions among the stakeholders (Larson & Larson, 2009).

Granted, there are always forces outside the control of those leading a project that will be distracting and affect the adoption of an idea, practice, or philosophy, especially one based on emerging technologies. Mumford (1934) termed this effect as “technics,” explaining there is a psychological construct to invention that is much larger than the technology itself. He regretted that critics of his “Technics and Civilization” book “overestimated the role of tools and machines… and overlooked the more passive, static, feminine aspects… the role of the internal transformer” (Mumford, 1959, p. 529). ICTs are merely reflections of humans who create and consume them. They are living and evolving according to human need.

Several innovations have been proposed to overcome researcher’s dissatisfaction with traditional publishing. How the diffusion of one of several ICTs will finally affect society is outside the control of any one person or project. The community of adopters determines how a technological innovation will be used. This behavior, in turn, influences agile modifications that typically improve the innovation or at least make it more fitting for the society using it. The final cycle of innovation is continuous modifications and improvements. This is explained in the forum innovation agility model. How the innovation adoption is used in action inspires a baseline reassessment of the original project and restarts the flow of innovation and diffusion with a newly stated common purpose and strategic plans to achieve outcomes.

The evolving model of scholarly communication was at the center of PubMed Commons messaging, research products, and forum commenting, as evidenced below.

> I have written a lot about how I think the biggest problem in science communication today is the disproportionate value we place on where papers are published when assessing the validity and import of a work of science, and the contribution of its authors. And I have argued that the best way to change this is to develop a robust system of post publication peer review (PPPR) , in which works are assessed continuously after they are published so that flaws can be identified and corrected and so that the most credit is reserved for works that withstand the test of time. (Eisen, co-innovator of PubMed Commons, 2013, para. 1)

> This [PubMed Commons] has the potential to greatly enhance what we call the most important peer review, which is the peer review after publication. The peer review before publication is very limited. Just a few people see it and editors. But afterwards, it’s the whole world. And this… it’s called PubMed Commons, we think, will be a really important edition to PubMed and certainly its use and its content is something that

> METRICS is going to be looking at very, very closely. (Goodman, early adopter of PubMed Commons, 2013, 1:30)
>
> Have to say as soon as I heard about PubMed Commons and used it, in fact as soon as I got the email I was intrigued. It was impossible not to think this is a significant moment in post peer review commentary. (Heneghan, pilot adopters, 2013, para. 2)

Innovators, organizers, and adopters of PubMed Commons were thrown off course by changes in the traditional publication process and lulled into believing that just because they built the forum, people would come. But something happened on the way to the forum… no one was keeping the boat afloat and headed toward the medal stand.

This dissertation has repeatedly highlighted how PubMed Commons innovators, organizers, sponsors, and adopters lost focus on introducing and diffusing an online forum to facilitate open discussions about published scientific literature. Instead, they became distracted by the possibilities of changes in scholarly communication made possible with new ICTs and a growing interest in online interactions.

When innovating and project planning large-scale online forums, I suggest using the forum innovation agility model I have introduced in this dissertation to keep the project focused on the stated outcome(s).

5.3 Limitations

This work is limited by the fact I was not able to member-check with Bastian or Vaught after the dissertation was completed. Although we connected via email with good intentions, our schedules to discuss my finished work was not possible. I would have welcomed the opportunity to know if project management strategies were being followed outside of the public eye. Following Creswell's (2007) advice, however, this dissertation was reviewed by faculty members on my committee and by my peers who have completed graduate studies in library and information sciences and knowledge and project management.

Related to this limitation, I purposefully designed data collection and analyses to be observational, gathering impressions of PubMed Commons as a PubMed user (e.g., medical librarian, patient) or potential adopter (e.g., forum commenter). Although my dissertation proposal and IRB approval included a survey of forum adopters, PubMed Commons was discontinued only days before my proposal defense, making it impractical to obtain perceptions about a group of individuals no longer part forum adopters.

In online research, two groups are difficult to access: lurkers and non-adopters. For the former, their participation is not visible, although they may benefit for reading the discourse. Non-adopters might have important insights about agile factors that organizers could revise to increase adoption or use To access these groups, forum organize could post a message on the forum inviting anonymous participation in an online survey. Lurker's input could be included in iteratively evaluating the forum and making agile adjustments to user experience functionality, design, and active participation. In the case of PubMed Commons, it would have been interesting to know how many adopters would have preferred to post anonymously, especially since data analysis showed that this issue was greatly important to forum adopters.

Since comments that violated forum guidelines were removed by moderators, it is impossible to measure what might have been considered an inflammatory.

5.4 Future Research

Applying the proposed forum innovation agility model to a similar forum, especially during the project planning phase, and then following up with a longitudinal investigation as a confirmatory study would offer the opportunity to test the model. Of interest would be comparing forum rate of growth and sustainability with PubMed Commons, especially given that recommendations presented in this dissertation (i.e., project management, user surveys, SNA measures, promotions, segmented commenter groups) were utilized. This research would also lend itself to evaluating user experience from the lurkers' point of view by soliciting real time feedback via an anonymous survey instrument linked from the library GUI. National Transportation Library (NTL) director Mary Moulton has shown an interest in my findings and is considering implementing a commenting system similar to PubMed Commons. Like PubMed, NTL is a global library with international users. Hence, there would be welcome similarities between forums.

PubMed Commons is a rich dataset that could be studied by researchers in a variety of fields. One area of interest is studying the effects of anonymity in online spaces. While this is a growing topic of research, few articles focus specifically on peer review (Bordignon, 2020; Teixeira da Silva, Al-Khatib, & Dobranszki, 2017). Based on the interest in this area and the number of passionate comments in PubMed Commons, it would be fascinating to undertake a follow up investigation comparing commenter identified comments on PubMed Commons with anonymous comments in PubPeer that were posted on the same PubMed-indexed articles in PubPeer. I quickly collected data to this effect on articles that received eight or more comments in PubMed Commons (see Table 5.4). This closed dataset provides unmatched opportunity.

A follow-up survey and interviews with PubMed Commons commenters could inspire continuing research about user experience, anonymity, and lessons learned from the forum. Outcomes could guide organizations who are interested in introducing, implementing, and managing an online forum to host scientific discussions about published articles.

Table 5.4

Comparison of Commeners on PubMed Commons and PubPeer

PMID	PubMed Commons Comments	PubPeer Comments as of 2/18/18	Anonymous Comments
26933091	33	51	18
26417050	26	359	233
25219520	22	–	–
24021304	16	–	–
26745426	14	14	0
25268438	13	17	1
25554788	12	47	34
27620683	12	–	–
28396415	12	30	0
28029926	11	–	–
28971835	11	8*	0
27518691	9	–	–
12053565	9	9	0
24733905	9	12	3
25415348	9	–	–
25739399	9	–	–
23363640	8	15	1
20143388	8	–	–
21334061	8	15	3
26129895	8	–	–
27693003	8	–	–
27934275	8	9	0
26524703	8	–	–
28074888	8	2	2
20877712	8	5	–

* Comments on PubMed Commons indicated as “deleted by users” are not posted on PubPeer.

Likewise, another avenue of research could investigate individuals who commented about PPPR on social media to advance knowledge about evolving perceptions related to peer review and scholarly communications.

Finally, a social network analysis of individuals who were PubMed Commons innovators, team leads, and organizers would demonstrate the strength of weak ties among what was obviously a who's who network of domestic and international academicians, government officials, and consumer organizations. Mapping their connections would add insights into the way such large-scale projects are initiated and managed. This information could inform about the rigor that goes into strategic planning by this type of high-profile network.

5.5 Conclusion

Was PubMed Commons merely "an experiment" (NCBI, 2018, para. 1)? Will it follow suit as another innovative idea introduced by the NLM that enjoyed only a short lifecycle under the stewardship of a large, federally funded agency before evolving and transforming into an idea that will be taken over by another enthusiastic, perhaps for-profit group at a later time? Are we waiting on necessary advancements in ICTs to move along the scientific revolution in scholarly communications (Kuhn, 1962)?

When reflecting on PubMed Commons and the case study presented herein, the most upsetting outcome is the missed opportunity. There is no comparable alternative for hosting an online forum for after-publication discussions about scientific literature. PubMed had a built-in audience and underlying mass communication platform. It also had an invested social system that was excited and vocal about their desire for changes to traditional publishing. For-profit organizations, including large publishing companies, continue to develop and promote systems that are forcing change in the way we do scholarly communication in science. Our government continues to implement policies requiring improvements in research transparency, research training for rigor and reproducibility, and open science and open access for federally funded research.

The big lesson learned from this retrospective explanatory case study research is that innovation adoption does not happen as if by magic. Just because you build it, they might not come. An innovation is only one required part for the diffusion of the innovation. A strategic plan with mindfulness on communication, time-to-adopt, and a social system to support adoption is critical for success. PubMed Commons innovators and organizers became distracted by what was outside their boat. Although they had a vision for how PubMed Commons would fit into the evolution of scholarly communication, they forgot that to become an adopted innovation in action, the forum had to remain sustainable. My goal for the forum innovation agility model is that it will guide organizations in strategically planning, implementing, and managing online forums for scientific discourse.

APPENDIX A

PUBMED COMMONS BLOG ARCHIVE

PubMed Commons editors published a WordPress blog to keep interested participants abreast of changes in the forum. The blog was active from November 26, 2013, through July 5, 2017. Several of the posts were cross-referenced in the NCBI Insights Blog, which aims to help readers understand and use resources of the NCBI at the U.S. National Library of Medicine. The blog archive below appears in its entirety; it is no longer accessible. Entries are provided in reverse chronological order with headlines and entry dates centered and highlighted. Some blog entries were accessible as of June 25, 2021 (e.g., https://ncbiinsights.ncbi.nlm.nih.gov/2013/10/23/joining-pubmed-commons-a-step-by-step-guide/ and https://ncbiinsights.ncbi.nlm.nih.gov/2013/11/01/early-developments-in-the-pubmed-commons-pilot/?relatedposts_exclude=315).

Collaborating to bring journal clubs to PubMed Commons: A librarian's perspective
Posted on July 5, 2017 by PubMed Commons Team

Journal clubs can be a great tool in graduate and medical education. They provide opportunities for students to practice important skills: literature searching, critical reading, scholarly debate, and in some cases, even writing. But are there ways to enrich the journal club experience? How can journal clubs become contributors to broader discourse? These questions intersect with traditional and evolving roles of librarians in higher education. Julie Hartwell shares how a collaboration with faculty on PubMed Commons got started and its initial impact.Before joining the Miller Nichols Library at the University of Missouri-Kansas City, Hartwell was in the A.R. Dykes Library at the University of Kansas Medical Center. When the PubMed Commons Journal Clubs pilot launched, Hartwell and her library colleagues were enthusiastic about bringing local journal club discussions to PubMed. So, she talked to the School of Nursing faculty about PubMed Commons Journal Clubs. Clinical Assistant Professor Chito Belchez shared her excitement about the idea.

Complementary missions

Evidence-based practice research is a core element of the Baccalaureate of Science in Nursing (BSN) program at the University of Kansas. Belchez was leading a course called "Nursing in an Evolving Healthcare System." For this and related courses, the journal club format offers flexibility to cover current developments in nursing practice. It also helps students develop the skills needed for critical literature review.

"They have to go through PubMed. They have to go through CINAHL [Cumulative Index to Nursing and Allied Health]. They have to go through all those databases to find an article outside their required readings to review," Hartwell describes. Students do their individual work outside class. In class, groups read and review the paper they've selected. Next, reviews go into Blackboard, an online learning management system (LMS). "They've been posting to discussion boards – creating a new thread, posting a review, and commenting on each others' work."

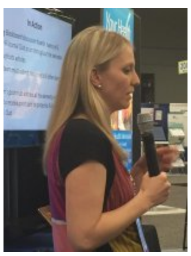
"Librarians have been trying to bring in this [Association of College and Research Libraries (ACRL)] standard of 'Scholarship as Conversation,'" Hartwell shares. The ACRL framework recognizes that research is not a one-and-done event. It casts scholarly discourse as an ongoing process, taking place in many venues. It suggests that, although systems may favor experts' voices, "novice learners" can contribute in meaningful ways.

But there are barriers to dissemination, even with online systems. Hartwell notes, "It's not brick and mortar so much as the walls of an LMS. You have all these great ideas… but they're hiding in the LMS where only your classmates can see them." It's one reason she was excited about the prospect of PubMed Commons Journal Clubs. It was a chance to expand the journal club's reach and to promote the principles of Scholarship as Conversation.

Building on frameworks

The University of Kansas School of Nursing Journal Club joined PubMed Commons in March 2015. As of June 2017, they've posted 23 comments.

Not every review written for the course makes it into PubMed Commons. "The students are given a rubric that's based kind of on the PRISMA [reporting] guidelines – what do you need to do to review an article," Hartwell explains. Although it was already in use, faculty began to think more about the rubric. They wanted to have clear guidelines for deciding which comments would appear in PubMed.



Hartwell talks about the journal club at the 2015 Annual Meeting of the Medical Library Association.

Along the way, Hartwell assisted with structure and support. She facilitated the PubMed Commons joining process. She was also able to get an inside look at what was being taught. "I was able to work with instructors on that rubric to make sure they were using the right terminology and providing students with the right resources," she notes. She was also able to see how students search for and select articles.

The addition of PubMed Commons posts seems to have provided a new incentive for students. They started following the rubric more closely. Faculty saw search strategies and writing improve. Hartwell comments, "The faculty will tell you that it's created a healthy kind of competition. 'We've got to do a really good review to get in PubMed.'" (You can read their comments here. They have a strong interest in nursing leadership and shared governance. They've also covered publications about workplace environment, quality improvement measures, and the impact of transnational migration on nursing workforces.)

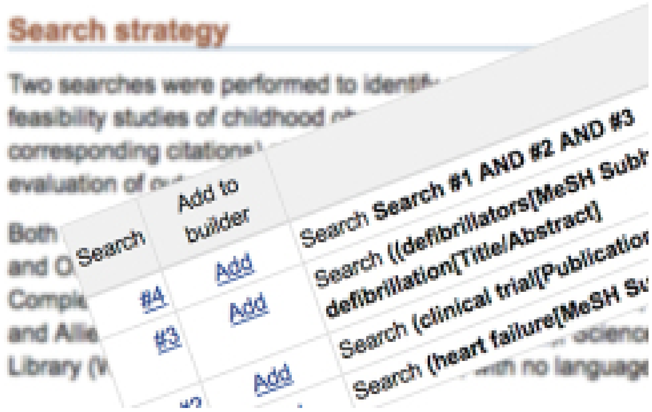
There are some unique advantages to librarians teaming up with journal club instructors. Hartwell shares, "What I found interesting was to… see what articles they're reviewing and how faculty communicate library resources to their students." Collaborations like this give librarians the chance to see what's being put into practice. And that means they can point faculty and students to up-to-date or alternate resources and provide tips for using databases like PubMed.

PubMed Commons also affords an opportunity to archive discussions and present them to a broader audience. "Without PubMed Commons, these good reviews and challenging questions would be lost in LMS. No one would ever see them again," Hartwell says. "Share. Don't leave these awesome enriching discussions hidden or, for face-to-face journal clubs, just lost. Preserve them."

Want to share and preserve your journal club's reviews on PubMed Commons? Learn more here.

Posted in Member Spotlight, On the Commons | Tagged Examples | Leave a reply

Critiquing systematic review search strategies on PubMed
Posted on May 28, 2017 by PubMed Commons Team



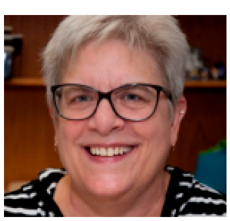
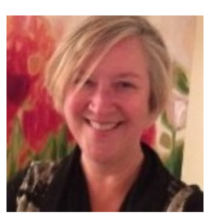
More than 1.1 million publications were indexed in PubMed in 2016, bringing the total number of PubMed records to more than 27 million. [See: https://www.nlm.nih.gov/bsd/licensee/2017_stats/2017_Totals.html]. It's no wonder that systematic reviews have become popular (currently there are more than 40,000 systematic reviews in PubMed Healthalone). Systematic reviews and related methods aim to pull together all relevant studies on a defined topic and synthesize the evidence to evaluate what's known. The approach has been used to inform clinical research and practice for decades, and its use is spreading.

As with any research, systematic reviews are only as good as their methods. A critical method here is literature searching. Some librarians and information specialists have taken to PubMed Commons to tackle issues surrounding the quality and efficacy of search strategies and their reporting. They also hope to raise awareness of librarians' expertise in this area. We interviewed 5 librarians to learn more about their perspectives and how they're using PubMed Commons.

Designing and reporting for reproducibility

Melissa Rethlefsen is deputy director of the Eccles Health Sciences Library at University of Utah and section director of the Systematic Review Core, which is integrated with the Center for Clinical and Translational Science. She has been investigating the quality of reported systematic review search strategies. She and colleagues at Mayo Clinic found that systematic reviews that included librarians as co-authors were more likely to meet standards such as those recommended by the Institute of Medicine. "It really does benefit you to have an information specialist or librarian on your team," Rethlefsen says.

"Just like any other type of research, your method should be described clearly enough that it can be r eproduced. We see so many systematic reviews that are published without this really critical information, and then it's really hard to assess their quality," Rethlefsen notes.

A number of journals have endorsed the use of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. Michelle Fiander, a systematic review librarian at the University of Utah, says, "PRISMA tells you what to say, what to report, the types of information that should be there. But they don't tell you how."

Mary Klem, a research and instruction librarian in the Health Sciences Library System at the University of Pittsburgh, has noticed the disconnect between statements concerning PRISMA and actual implementation. "In the article I commented on, the authors' primary rationale for completing their review was that a prior review on the topic had not used a systematic or well-defined search strategy," she shares. "I thought it was awesome that someone had critiqued a review like that! So I was disappointed to see that the documentation and searches in this new improved review weren't thorough or comprehensive, and felt like I needed to note that."

Putting expertise forward

Some librarians have used PubMed Commons because it's visible, it's fast, and they see potential for the impact to extend beyond a single publication.

"In a practical sense, using PubMed Commons seemed to be a better choice than writing a letter to the editor because of its immediacy and visibility," Donna Berryman, the director of the Miner Libraries in the Institute for Innovative Education at the University of Rochester Medical Center, says. "Many, many people will find the article I commented on by doing a search in PubMed. If they look at the record for the article, my comment will be there. I'm not sure how many people would even look at a letter to the editor. In addition, there's always a chance the letter won't get published, and, if it does, there's generally a long gap between when an article appears in a journal and when the letter to the editor might appear. All of those things argue against visibility. So, PubMed Commons gives my words visibility and immediacy."

Wichor Bramer also favors the transparency and timeliness of PubMed Commons. He is a biomedical information specialist at Erasmus University Medical Center in the Netherlands, where he's also currently working on his PhD on search methodologies for systematic reviews. He shares, "My last comment was on the details of a search strategy. Julie Glanville, who's a famous searcher for reviews, responded to that, so you can communicate publicly with the authors." Author responses can create a "vivid discussion that's available for anyone to see."

Bramer is also finding value as an author himself. He notes that he's used comments on his first article comparing PubMed and Google Scholar to change the way he did some things for his second article. "The comments that we get help me create better articles in the future."

For Fiander, PubMed Commons offers an opportunity to "get my voice out there and point out things. Maybe it will end up stirring some better standards among journal editors. If you have a paper and you're indexed in there, you can comment. It's easier than writing a letter to the editor. I think the freedom of it is good."

Commenting with care

But freedom doesn't mean off the cuff for these commenters. "I tend to read my comment, be careful that I'm being accurate, that I'm not overstating or saying something that's inaccurate," Fiander notes.

Berryman has commented once but suspects she will comment again. "PubMed Commons strikes me as a place to have reasoned, deliberate comments. It's not like commenting on Facebook or Twitter. So, one thing I always think about is whether I can write my comment in a way that is constructive and will add to the body of knowledge – and that takes both thought and time."

"I see it as post-publication peer review," Bramer says. "I first create it. I don't post it immediately. I put it away for maybe a day and look at it the next day and see different things, see if I can improve some things."

Rethlefsen understands that commenting on PubMed, especially the first time, isn't necessarily easy. "It's not really a space where librarians had actively engaged before. Irreproducible search strategies were always a thing that librarians talked to each other about." But she had concerns about what looked like, on the face of it, an excellent search strategy that she couldn't reproduce. So she decided to go to PubMed Commons. "I worked on it, I deleted it, I re-wrote it, and I deleted it. And finally, I pushed the publish button. It was intimidating because I'd never done it before. But once I got the reaction that I did [from colleagues and the librarian community], it became really clear to me that this was a more important space than I'd thought before."

Joining in

Feeling inspired? If you have a publication indexed in PubMed, then you're eligible to join PubMed Commons and start commenting! Learn more about getting started with PubMed Commons.

Posted in Member Spotlight, On the Commons | Tagged Examples | Leave a reply

Authors alerting readers via PubMed Commons
Posted on April 6, 2017 by PubMed Commons Team

Journals can issue correction and errata notices to notify readers of errors and, as necessary, revise text and data in publications. Yet these processes can take time. Authors sometimes encounter obstacles to publishing corrections. Some authors use PubMed Commons to alert readers to issues or to refine language and interpretations. Correcting the record via journal notices is important, and it's great to see authors add speed and transparency with post-publication updates.

Earlier this year, Garret Stuber commented on a publication on hormonal control in social reward. In the days after it appeared in press, some errors came to his attention, which he was working to address through a formal correction notice. Stuber told Retraction Watch that, in the meantime, he commented "in an effort for immediate notice and transparency to what occurred."

Sometimes an error may significantly change the results of a study. Stefan Hofmann commented on a meta-analysis he co-authored, examining the effects of oxytocin on an array of psychiatric symptoms. However, readers raised some points, prompting another look at the data. Errors were made in specifying the direction of outcomes investigated. Hofmann reported the re-calculated effect sizes and indicated that the article is being retracted.

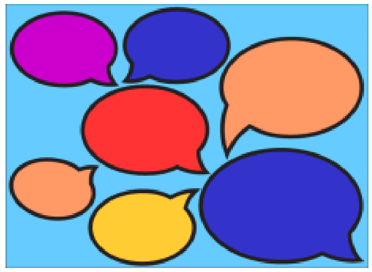
Here are some more authors setting this great example:

- Ahmet Selçuk Can, on behalf of authors, posted a correction for text and a table where a journal is no longer active.
- Wichor Bramer clarified a step in de-duplicating database search results for systematic reviews in reference management software.
- Michael Hoffman noted a minor typographical error in the online methods section for his first-authored publication on pattern discovery in human chromatin structure.
- Jonathan Eisen highlighted a "sentence in the paper that could be worded more carefully" concerning the draft genome of an actinobacteria.
- Following widespread misinterpretation of his findings about false-positive rates in function magnetic resonance imaging (fMRI), Thomas Nichols posted a revised significance statement. (You can read more about discussion surrounding that publication and comment here.)
- Kevin Hall linked to and posted a published correction for a co-authored publication on metabolic adaptation.
- Alexander Tsai corrected values reported in the abstract for his systematic review and meta-analysis of depression assessment among persons with HIV in sub-Saharan Africa.
- Edward Berry noted an incorrect metal assignment in protein crystal structure, pointing readers to a source for additional information.
- John Denning posted a correction concerning cutoffs used in a test of memory malingering, which was later published in the journal.

Journal corrections revise the version of record for a publication, and PubMed Commons does not replace that. But it does offer another way for authors to provide clarifications, point to interim and published corrections, and alert readers to errors quickly. And it's good to see authors taking advantage of PubMed Commons to pass that information along to the community.

Posted in On the Commons | Leave a reply

PubMed comments & their continuing conversations
Posted on November 21, 2016 by PubMed Commons Team



We have many options for communication. We can choose platforms that fit our style, approach, and time constraints. From pop culture to current events, information and opinions are shared and discussed across multiple channels. And scientific publications are no exception.

PubMed Commons was established to enable commenting in PubMed, the largest biomedical literature database. In the past year, commenters posted to more than 1,400 publications. Of those publications, 80% have a single comment today, and 12% have comments from multiple members. The conversation carries forward in other venues.

Sometimes comments pull in discussion from other locations or spark exchanges elsewhere. Here are a few examples where social media prompted PubMed Commons posts or continued the commentary on publications.

Debating disease association

On June 3, 2016, Daniel MacArthur took to Twitter to express his skepticism of a report describing an association between a gene mutation and familial multiple sclerosis published in the journal *Neuron*. His critique stirred a bit of interest. A few days later, he posted a comment, co-written with Eric Minikel, to PubMed Commons. MacArthur and Minikel highlighted, "Enrichment in cases over controls is one important criterion for establishing pathogenicity of sequence variants." The comment prompted more discussion on Twitter.

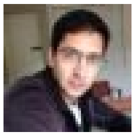
 Daniel MacArthur@dgmacarthur
7:42 AM - Jun 13, 2016
PubMed Commons comment by @cureffi and me about that "new MS gene" paper:
http://www.ncbi.nlm.nih.gov/pubmed/27253448#cm27253448_16159 … Basically, there's no way this is real.
1111 Replies
6969 Retweets
6060 likes

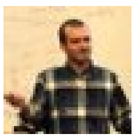
Twitter Ads info and privacy

Over the following days, author Carles Vilariño-Güell responded, and MacArthur and Minikel replied. Shortly Chris Cotsapas posted a comment on behalf of the International Multiple Sclerosis Genetics Consortium. His comment summarized an attempt to validate the findings, linking to results posted in a bioRχiv preprint. With Simon Heath, Daniel Weeks noted further concerns in an August comment on the journal's website, which he linked from PubMed Commons.

As the critiques unfolded, some readers commented on blog posts highlighting the results, (such as here and here) to point to the comments on PubMed. In September, STAT published a story reviewing the concerns that had been raised on PubMed Commons and elsewhere. In October, *Neuron* published letters from the International Multiple Sclerosis Genetics Consortium and Minikel and MacArthur, as well as a response from the authors. In an accompanying editorial note, the editors remark that the peer-reviewed letters offer "an important complement to other forms of commentary" including social media, PubMed Commons, and the journal's online comments section.

Comments also sparked discussion of topics beyond the specific gene variant in question.

 Lior Pachter@lpachter
Jun 18, 2016
An interesting exchange between @dgmacarthur, @cureffi and Charles Vilarino-Guell on @PubMedCommmons http://ncbi.nlm.gov.pubmed/27253448#cm27253448_16258

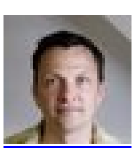
 Joe Pickrell@joe_pickrell
10:23 AM - Jun 18, 2016
I suspect analysis of the LRRK2 variant would show it is massively enriched in Parkinson's cases over ExAC
22 Replies
11 Retweet
55 likes

 Divad Retsop@DivadRetsop
Replying to @dgmacarthur
5:18 PM - Jun 13, 2016
@leonidkruglyak @cureffi Wonder how variants in oligogenic Bardet-Biedl syndrome-type inheritance will fare in ExAC comparisons
11 Replies

Self-correcting statements

In July 2016, a publication co-authored by Thomas Nichols reported on an artifact that might give rise to high false-positive rates in functional magnetic resonance imaging (fMRI) analyses. Across blogs and professional publications, there was consideration of what the findings meant for neuroscience research. But some in the community thought results were being too broadly extrapolated to all fMRI studies, not just the specific issues examined.

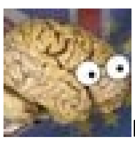
As the publication was discussed online, the authors recognized that some wording was being interpreted in ways they had not anticipated. So they asked to publish an erratum. That was initially rejected by the journal, since there was no change to the results or conclusions. Nichols published the note on his blog. Following an exchange on Twitter, he subsequently posted a comment on PubMed Commons to make a more circumspect significance statement.

 Marcus Munafo@MarcusMunafo
Add it as a comment on @PubMedCommons – that way anyone who sees the abstract on PubMed will see the comment.

 Thomas Nichols@ten_photos
@MarcusMunafo Brilliant idea! Pubmed Commons comment now live:
http://www.ncbi.nlm.nih.gov/pubmed/27357684

1:37 PM - Jul 14, 2016
11 Reply
1111 Retweets
1818 likes

At least one blogger updated a post to reflect the authors' statement.

Neuroskeptic@Neuro_Skeptic
Replying to @Neuro_Skeptic @ten_photos and @MarcusMunafo
I have updated my post http://blogs.discovermagazine.com/neuroskeptic/2016/07/07/false-positive-fmri-mainstream/#.V4is1BLSM4k …
3:29 AM - Jul 15, 2016

False-Positive fMRI Hits the Mainstream – Neuroskeptic
A new paper in PNAS has ade waves. The article, called Cluter failure: Why fMRI inferences for spatial extent have inflated false-popsitive rates, comes from Swedish neirosceintists Anders Eklund,…
22 likes

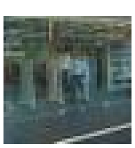
Although the journal ultimately published a correction a month later, PubMed Commons enabled authors to rapidly communicate a reframed interpretation of their work.

Replicating and reviewing search strategies

Comments can initiate discussion of specific results and interpretations. But they can also serve as a jumping off point to evaluate approaches and highlight practices.

Literature search strategies lie at the core of systematic reviews and meta-analyses. Melissa Rethlefsen posted a comment describing an attempt to replicate the search strategy reported in a meta-analysis. She noted key missing information such as date ranges. She concluded: "This study highlights the need for more accurate and comprehensive reporting needed for search strategies in systematic reviews and other literature search-based research syntheses, and the need for better peer review of search strategies by information specialists/medical librarians."

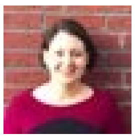
One library used this example to encourage the use of structured reporting guidelines for systematic reviews and meta-analyses.

 Highland Health@HHSLNew

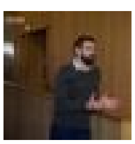
Move from good to excellent - conform to PRISMA - use the libray service - see comments on search strategy at - http://tinyurl.com/zht54tn
7:50 AM - Apr 6, 2016

Another library used the comment to illustrate the importance of reviewing search strategies. And medical librarians and researchers chimed in on Twitter.

PubMed Commons@PubMedCommons
Replicating literature searches: M Rethlefsen/@mlrethlefsen highlights impact of search strategy reporting. http://1.usa.gov/1QSeUiJ


StephHKinsler@BookTechno
@PubMedCommons @mlrethlefsen Wowza, that's great work. I am a new med lib gig but I would think this is a significant issue...
2:20 PM - Apr 6, 2016
11 Reply
11 like


Isla Kuhn@ilk21
@brie_mcc @krafty http://www.ncbi.nlm.nih.gov/pubmed/26009228#cm26009228_14545 … raises good point about #medlibs (lack of) involvement in peer review process around reporting
10:27 AM - Apr 6, 2016
11 like


Ian Lahart PhD@IMLahart
Importance of accurately describing search strategy in systematic reviews highlighted here

👏 http://www.ncbi.nlm.nih.gov/pubmed/26009228#cm26009228_14545 … pic.twitter.com/sj2xQM3s1F
2:19 AM - Apr 7, 2016
11 Reply
88 Retweets
1010 likes

Extending the reach of scientific discourse

As you browse the web, you might just run across a mention of a comment on PubMed. Blog authors and readers might mention comments, as they have about a genetic variant associated with body mass index , 'bad luck' and cancer, or the occurrence of amphetamines in water systems. They might even appear in the references list, such as a roundup of publications on cancer risk or a look at psychological debriefing after traumatic events. Perhaps the most talked-about comments were those from Jennifer Doudna and Emmanuelle Charpentier on a perspective of the history of CRISPR. The comments were shared and discussed on social media. They were also mentioned in several blog posts and articles, including ones from news outlets such as the *Washington Post.*

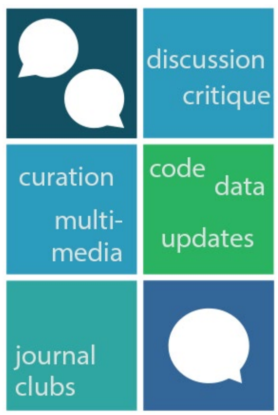
PubMed Commons offers a key place to quickly anchor critical points for future readers to see—in one of biomedical science's most heavily used resources. On a typical day in 2016, 1.6 million users ran 2.5 million web searches on PubMed and viewed more 8 million records.

Through October 31, 2016, PubMed Commons had 10,632 members. They've posted 5,739 comments to 4,595 publications. Want to join in? Check out our Get Started page for more information!

____

The PubMed Commons Team
Posted in On the Commons | Tagged Examples, Popular | 1 Reply

Commenting on PubMed: A Successful Pilot
Posted on December 17, 2015 by PubMed Commons Team



We are pleased to announce that PubMed Commons is here to stay! After developing and piloting the core commenting system for PubMed, a pilot of journal clubs was added. And we have completed a major internal evaluation of the use of the Commons. We aim to publish that soon, so stay tuned to this blog or Twitter for news on that.

PubMed Commons provides a forum for scientific discourse that is integrated with PubMed, a major database of citations to the biomedical literature. Any author of a publication in PubMed is eligible to join and post comments to any citation.

More than 9,500 authors have joined PubMed Commons – and they have posted over 4,000 comments to more than 3,300 publications, mostly on recent publications. Commenting has plateaued, so the volume is low. But the value of comments has remained high. And comments often attract a lot of attention.

About half the comments are on clinical or health-related publications. Members have been using PubMed Commons to:

Update and expand the public record, for instance by pointing to new data, relevant publications, or alternative interpretations

Note corrections and retractions to publications

Post discussion and critique, either directly or via links to blog posts and other platforms

Provide links to datasets, code, or publicly accessible versions of publications

Call attention to issues affecting reproducibility, such as cell line misidentification

Authors posting to their own publications contribute about one in five comments. About one-third of these have been replies to questions or discussion from others. Since the PubMed Commons Team began notifying authors of comments on their publications, the proportion of comments with author replies has increased. However, the rate of reply remains below 10%. We will keep working on ways to encourage more author response.

Just a year ago, we introduced a new mechanism to capture the synthesis of journal club discussions of scientific publications. PubMed Commons Journal Clubs have full commenting privileges and profile pages to provide background information about the club. To date, 20 journal clubs have joined. These institutional, virtual, and hybrid journal clubs represent a range of clinical and biomedical disciplines. They have become a critical and vibrant part of PubMed, and we are planning more support for this initiative.

PubMed will shortly turn 20. It has become a major resource for finding biomedical and health-related literature. There are now more than 25 million citations. And there were more than 2.7 billion searches in the last year – that's more than 7 million searches a day.

That means that comments have a large potential audience, and the interest in them is growing. Visits to the PubMed Commons homepage have nearly doubled, from 1.2 million in the first half of 2014 to 2.3 million in the first half of 2015.

We believe the commenting function addresses a critical need, for PubMed and for the development of biomedical research. So a big "thank you" from us to everyone who has contributed their time and energy to supporting the Commons and commenting at PubMed.

Just because the pilot has ended, doesn't mean PubMed Commons will stop evolving. With the pilot over, we're working on an application program interface (API) that will enable hosting of PubMed comments on third-party sites. And other new features are in the pipeline. Meanwhile, anyone can submit suggestions and feedback by using the "Write to the Help Desk" link at the bottom of NCBI pages.

Ready to get involved? Visit our Getting Started page to learn more about how to join and participate in PubMed Commons – or start here if you would like your Journal Club to join in.

____

The PubMed Commons Team

Posted in Commons News | Tagged News | 3 Replies

Signposts from research to resources

Posted on May 6, 2015 by PubMed Commons Team

From repositories to blogs, the web has expanded means to share information and resources widely. Access to data and code enables other researchers to check published analyses and undertake new ones. Having another way to look at results can help people connect with them and deepen understanding. PubMed Commons members are tying these pieces back to publications by adding external links to PubMed records.

Tagging inputs

High-throughput assays generate heaps of data, which can require custom software tools to process and analyze. Some authors are annotating current locations and updates for data and code via PubMed Commons.

Proteomics studies approach a wide range of questions about proteins and pathways, often with mass spectrometry data at the core. Author David Simpson provides the identifier and URL to access the dataset for a recent publication. Attila Csordas has also connected several proteomics articles to deposited data.

Patrick Schloss and colleagues published an approach for characterizing microbiomes using a particular high-throughput sequencing platform. He links to "a fully executable version" of his paper. The repository includes the R code, as well as raw and processed data, so that users can reproduce results in the publication.

With the end of Google Code on the horizon, researchers are moving projects to new locations. Pedro Mendes has migrated code for a tool used in modeling of biochemical networks to GitHub. He's added a comment to point to the code's new home.

Sometimes authors will update code and append new options. Ross Lazarus summarizes features added to a toolkit for high-throughput biology workflow software. He also includes a link to the new version.

Adding dimensions

Three-dimensional structures of biological molecules can offer useful insight into how proteins function. But as figures in papers, structures can fall flat. Some are using PubMed Commons to restore depth.

Michael Cianfrocco and colleagues solved the structure of a transcription factor complex bound to DNA. He provides a link to FigShare where users can download files for a visualization program. They can then dive into the structure and even create their own figures.

Sandra Porter links to a blog post, where she writes, “One of the most amazing things, to me at least, is how spider silk changes from a liquid form, inside the spider, to a solid, strong material that we see in their webs and other constructions.” She shows readers how to use their tablets to explore the structure and properties of a protein in spider silk that permit this change.

Mary Mangan offers a resource for a literal hands-on approach. She used data from an X-ray crystal structure to create a 3D-printable model of γ-hemolysin, a pore forming protein from *Staphylococcus aureus.* She points readers to the model on the NIH 3D Print Exchange.

Have something you want to add to a publication? Any author of a PubMed-indexed publication is eligible to join PubMed Commons. Learn how! And check out more examples of how PubMed Commons is being put to use.

____

The PubMed Commons Team

Posted in On the Commons | Tagged Authors, Examples | Leave a reply

Introducing PubMed Commons Journal Clubs

Posted on December 17, 2014 by PubMed Commons Team

Around the world, the journal club is a cornerstone engagement with the scholarly literature. Whether in face-to-face meetings or on social media platforms, researchers, physicians, and trainees gather to debate and converse about publications. Participants share their views on methods and interpretations of results. They discuss how publications fit into a broader context or might inform their own research or practice.

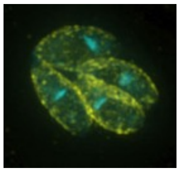
In short, the journal club can represent a major intellectual investment – and a long-standing form of post-publication evaluation.

Yet often, the analyses and ideas don’t travel far beyond core participants. Digital records and virtual journal clubs can help deliver the discourse to others. Still, wouldn’t it be fantastic if more of us could see what these groups have to say?

Today we’re excited to announce the launch of PubMed Commons Journal Clubs. These accounts will allow groups to establish their own identity on PubMed Commons. Journal clubs will be able to share key points, questions, and summaries from their discussions – right below citations in PubMed.

Bringing local discussion to the global Commons

Gary Ward is a professor in the Department of Microbiology and Molecular Genetics at the University of Vermont. (He is also a member of the external working group providing feedback on PubMed Commons.) His lab studies *Toxoplasma gondii*, a protozoan parasite. It’s widespread among humans and other mammals and can cause serious illness for those who are pregnant or have weakened immune systems.



UVM Toxo Journal Club covers work on parasites like *Toxoplasma gondii.*

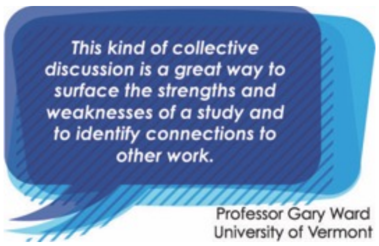
Ward also facilitates the University of Vermont (UVM) Toxoplasma Journal Club, a group of grad students, postdocs, technicians and faculty who do research on *T. gondii*. “We try to review both classic papers (why is this a classic in our field?) and very recent findings in the world of parasite cell biology.”

“We each take turns picking a paper and leading the discussion,” he explains. Last year, the group added a new step. “Immediately after the journal club, the discussion leader is responsible for drafting a PubMed Commons comment that summarizes the key points of the discussion. The comment is revised based on feedback from the group and then posted.”

Ward notes the direct benefit of this process for participants. “Having to summarize our meeting in the form of a comment forces us to distill the many things that were discussed into the two or three most important points. The ability to focus one’s critique/comments in this way is a great skill for grad students and postdocs to learn, and for the rest of us to practice.”

He also thinks that journal clubs have something more to offer to the scientific community at large. “Other than the journal club setting, how often does a paper get read critically from beginning to end by 10-12 informed readers who then discuss it at length as a group? This kind of collective discussion is a great way to surface the strengths and weaknesses of a study and to identify connections to other work.”

“Posting journal club comments in PubMed Commons adds depth to the literature and may give the reader a different perspective on the work,” Ward explains. “They will be particularly useful when they stimulate the authors to engage in a PubMed Commons dialog. If our journal club had a particular question about the paper, it is likely that other readers will as well.”



The UVM Toxoplasma Journal Club has a great example of just how that can happen.

Expediting lab-to-lab communications

Three thousand miles away from Burlington, Vermont, Markus Meissner’s group at the University of Glasgow had worked out a method to target genes in *T. gondii* for conditional deletion. They applied the approach to look at how the parasite infected host cells. Meissner’s group found that actin was essential to *T. gondii* survival – but not because the parasites couldn’t invade host cells. Rather, they argued, the parasites die because they lose a specialized part of the cell called the apicoplast.

“In our discussion of this paper,” Ward notes, “a new graduate student in the group suggested a great idea on how to test this hypothesis.”

The apicoplast is essential for survival of *Plasmodium falciparum,* the parasite that causes malaria. However, blood-stage *P. falciparum* can live without an apicoplast if supplied with isopentenyl pyrophosphate (IPP), which is normally produced in the apicoplast.

The journal club asked in their comment: Could *T. gondii* lacking actin survive if given IPP?

Meissner replied. His lab had considered the experiment but scrapped the idea after learning from other experts that IPP treatment doesn’t have the same effect in *T. gondii* as it does *P. falciparum.*

This instance illustrates how PubMed Commons can initiate useful exchanges. “Now anyone wondering if IPP rescues an apicoplast defect in *T. gondii* can discover that it doesn’t,” says Ward. “That information had not previously been captured, but now it is in the form of a PubMed Commons comment.”

Calling journal clubs to join the discourse

NephJC brings discussions from the nephrology (& related specialties) Twitter community.

With PubMed Commons Journal Clubs, we’re hoping to see groups and individuals engaging on PubMed Commons and beyond. We’re pleased to welcome the UVM Toxo Journal Club, NephJC, and CREBP Journal Club as our first PubMed Commons Journal Clubs.

To encourage connections, PubMed Commons Journal Clubs will have profile pages on PubMed Commons. These pages will provide descriptions of the groups and ways to connect with them outside PubMed Commons (click the Journal Club images in this post to see their pages). We’re also starting a Facebook page to offer a space for group members to start sharing their ideas (link coming soon). We’ll be exploring other ways to help groups network, as we build and develop the PubMed Commons Journal Clubs community.

CREBP Journal Club at Bond University looks at the gaps between evidence and current clinical practice.

PubMed Commons Journal Club accounts are currently open to journal clubs discussing literature for research, graduate and postgraduate education, or continuing professional education. Applications will need to be supported by PubMed Commons members who participate in the group's discussions. For more information or to apply for a Journal Club account, email pubmed.commons@ncbi.nlm.nih.gov.

____

The PubMed Commons Team

Posted in Commons News | Tagged News, Update | 1 Reply

Comment search and alert: A PubMed Commons guide
Posted on December 13, 2013 by NCBI Staff

Some authors are now adding comments to PubMed records in the pilot PubMed Commons project. Soon, these comments will be visible.

How can you find these needles in the giant PubMed haystack? How can you know if someone has commented on your publication, or joined a discussion on an article you're interested in? We'll show you how to find articles with comments first – and then how to use these searches to get alerts on new comments.

This is the key piece of PubMed search language you need:

| AND has_user_comments[filter]

Put whatever you are interested in front of that, and only those publications in PubMed with a comment will appear. You can use this filter to find articles on particular subjects, names, journals and much more.

To find out if there are comments on a particular article:

PMID is the acronym for a record's ID in PubMed. You can see it at the end of the abstract view – PMID: 11572773. Here's how you use it to find out if it has a comment:

| 11572773 [pmid] AND has_user_comments [filter]

If there is no comment, the search will come up empty.

To find out if there are comments on articles by a particular author:

We recommend this technique, with the author's last name followed by initials, without punctuation:

| Chimenko I [author] AND has_user_comments [filter]

You can shorten [author] to [au]. This technique also works for full names for many publications since 2002, like this:

| Chimenko, Ingrid [author] AND has_user_comments [filter]

If you have a unique author identifier, it will only work for the articles where the publisher has included the number in the PubMed data.

We're working on ways to make your own articles quicker to target. In the meantime, you could check out the video tutorial on PubMed searching by author.

PubMed has many other pieces of search language you can use to target other things you want to find. There is a list here in PubMed Help.

Keep PubMed on the alert for new comments for you

So how can you use these searches to get alerts for new comments? For that, you need a My NCBI account. My NCBI is free and open to anyone. When you use it to set up alerts for searches, you will get an automatic email alert from PubMed when the search finds a new comment. Check out NCBI's how-to guide and PubMed will be on the alert for you.

The PubMed Commons team

Coming next on PubMed Commons blog: how authors are using the Commons to expand, update and correct the records of their work.

More information.

Setting up automatic NCBI searches and new record alerts

How to join PubMed Commons

(Note – if you are an author of a publication in PubMed Commons, your email may be in the list explained in this post)

Posted in Tips | Tagged Alerts, Authors, Search | Leave a reply

PubMed Commons going public soon
Posted on November 26, 2013 by NCBI Staff

It's been a month since the beta launch of PubMed Commons, the pilot system that enables authors' discussion and sharing of information about publications in PubMed.

The first public version of the PubMed Commons pilot will be released in the coming weeks. All users of PubMed will be able to see and cite comments.

We're grateful to the hundreds of you who joined the closed phase of testing – especially for your patience with the inevitable bugs in a beta system. Your activity and feedback have made the system better in several ways:

- There will be a simplified way for eligible authors to join – including all those with current author email addresses in PubMed and PubMed Central;
- A permanent citable link will be available;
- We have increased the space in individual comment boxes (up to 8000 characters), and the new release will have warnings if you're getting close;
- Article helpfulness ratings are influencing the comment stream on the home page;
- New specific guidelines have been released to address concerns reported by members.

More features are in the pipeline. There will be an increased use of data from the helpfulness ratings to make the display of comments more helpful – and we will support sharing on social media. Also in development is an application programing interface (API) to integrate comments from PubMed Commons into other websites.

Exploring options to ensure a vibrant and useful forum for discussion of scientific publications will be a key focus of the next stage of the pilot. We will be exploring ways to expand people's access to commenting and rating helpfulness, for example through group accounts. Enhancing the value of PubMed for users is critical to the success of PubMed Commons, and we are relying on the community to help shape the conduct and system it wants to see.

We are establishing a working group to advise us during the next stages of the pilot and its evaluation. And we look forward to community discussion, too.

The upcoming release marks the start of evaluation of PubMed Commons. Evaluation results will be considered at 3 months and 6 months, with the final report on the pilot anticipated after 9 months. Three key areas will be our focus: uptake and reputation, quality and impact of comments and discussion, and sustainability.

We will be blogging more about comments being made in the Commons, what we're learning, and explaining more about aspects of the system. We will be trying out a Twitter chat too, so keep your eye out on @PubMedCommons for the announcement.

Thanks again to everyone who has contributed to PubMed Commons and the discussion about it. We look forward to an even wider discussion soon. Stay tuned to this blog or @PubMedCommons for news of the Commons going public.

____

The PubMed Commons team

More information

How to join PubMed Commons *(Note – if you are an author of a publication in PubMed Commons, your email may be in the list explained in this post.*

Posted in Commons News | Tagged News | 5 Replies

Welcome!
Posted on November 14, 2013 by NCBI Staff

Welcome to the new PubMed Commons blog!
Soon we will begin posting more information about the PubMed Commons project, including new and anticipated features, notable comments and topics of special interest.

____

Posted in Uncategorized | Leave a reply

ABOUT PUBMED COMMONS
PubMed Commons is a pilot commenting system for authors in PubMed®. PubMed is the U.S. National Library of Medicine's database of the biomedical literature.
PUBMED COMMONS LINKS
PubMed Commons Home (https://www.ncbi.nlm.nih.gov/pubmedcommons/)
PubMed Commons on Twitter (https://twitter.com/pubmedcommons/)
National Center for Biotechnology Information (NCBI)
RECENT POSTS
Collaborating to bring journal clubs to PubMed Commons: A librarian's perspective
Critiquing systematic review search strategies on PubMed
Authors alerting readers via PubMed Commons
PubMed comments & their continuing conversations
Commenting on PubMed: A Successful Pilot
Blog at WordPress.com

APPENDIX B

PUBMED COMMONS GUIDELINES

PubMed Commons was promoted as a system that enabled authors to share opinions and information about scientific publications. All authors of publications with PubMed records were eligible to become members and comment on any publication in PubMed. They were also privileged to invite other eligible authors to join. Journal Clubs could apply for membership with an individual PubMed Commons members serving as a guarantor. The official language of PubMed Commons is English.

Members were told they play a pivotal role in ensuring that PubMed Commons remains a forum for open constructive criticism and discussion of scientific issues, by reporting concerns and rating the helpfulness of comments. Members making substantive criticisms of a publication were encouraged to let the publication's authors know there is a comment. An "invite an author to comment" facility is provided at each PubMed entry.

By joining PubMed Commons, members agreed to the following:

- Establish a single individual account with their real name (no pseudonyms or anonymous accounts are allowed)
- Follow the current guidelines when they use PubMed Commons
- Grant other users a worldwide, royalty-free, non-exclusive, perpetual license under the Creative Commons Attribution 3.0 United States License
- Invite only eligible individuals to join PubMed Commons
- Disclose potential conflicts of interest
- Make comments that are directly relevant to the particular work in PubMed on which they are commenting
- Not use PubMed Commons to spam or systematically promote a product, position or the members' own publications or to target others

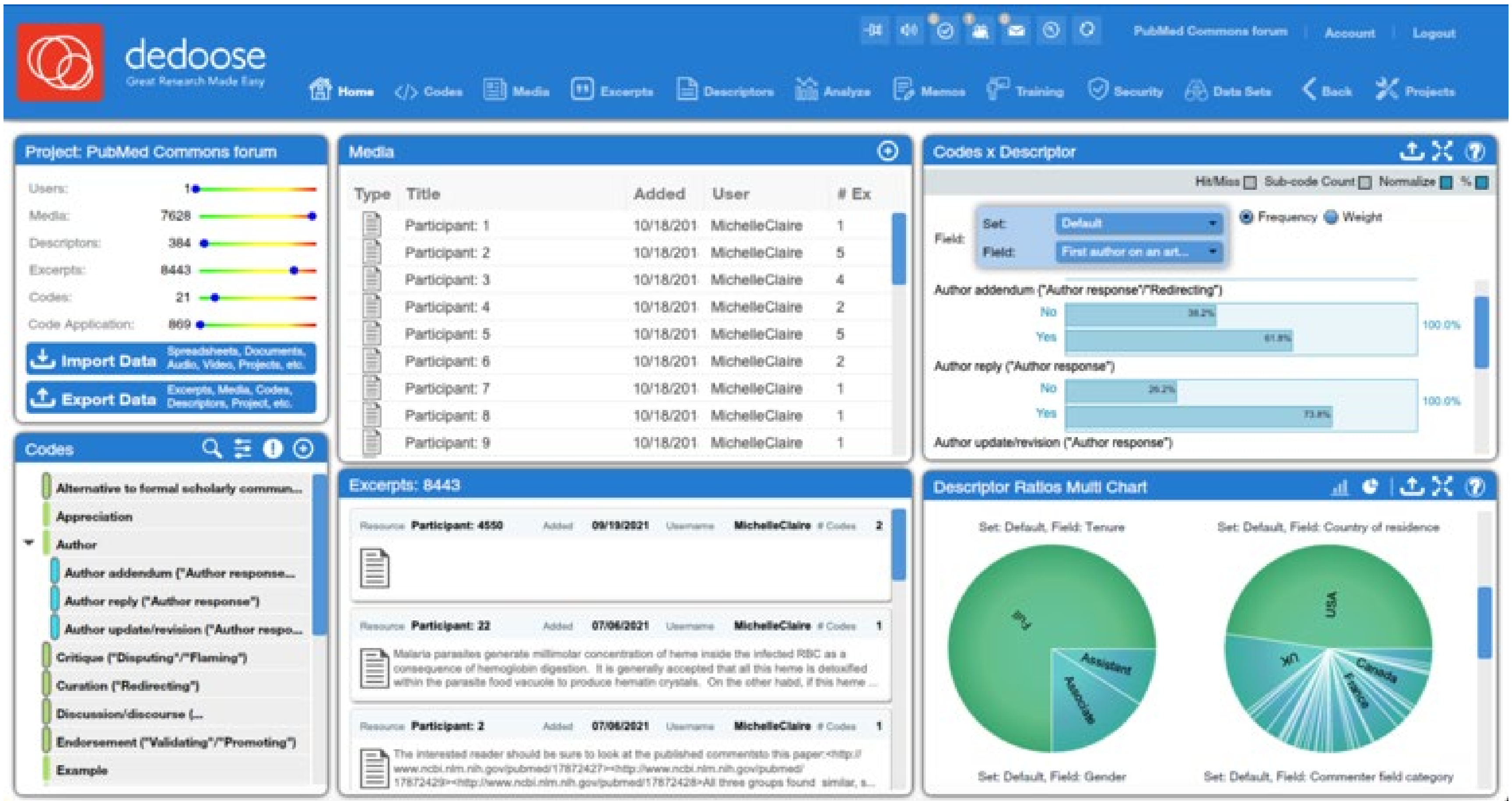
And comments should not contain:

- Discriminatory, racist, offensive, inflammatory, unlawful, or derogatory language
- Partisan political views
- Plagiarized content
- Descriptions or content of unpublished work by others without permission
- Allegations of misconduct on the part of authors, reviewers, editors and publishers
- Speculation about the motivations of authors, reviewers, editors and publishers

APPENDIX C

DEDOOSE DATABASE

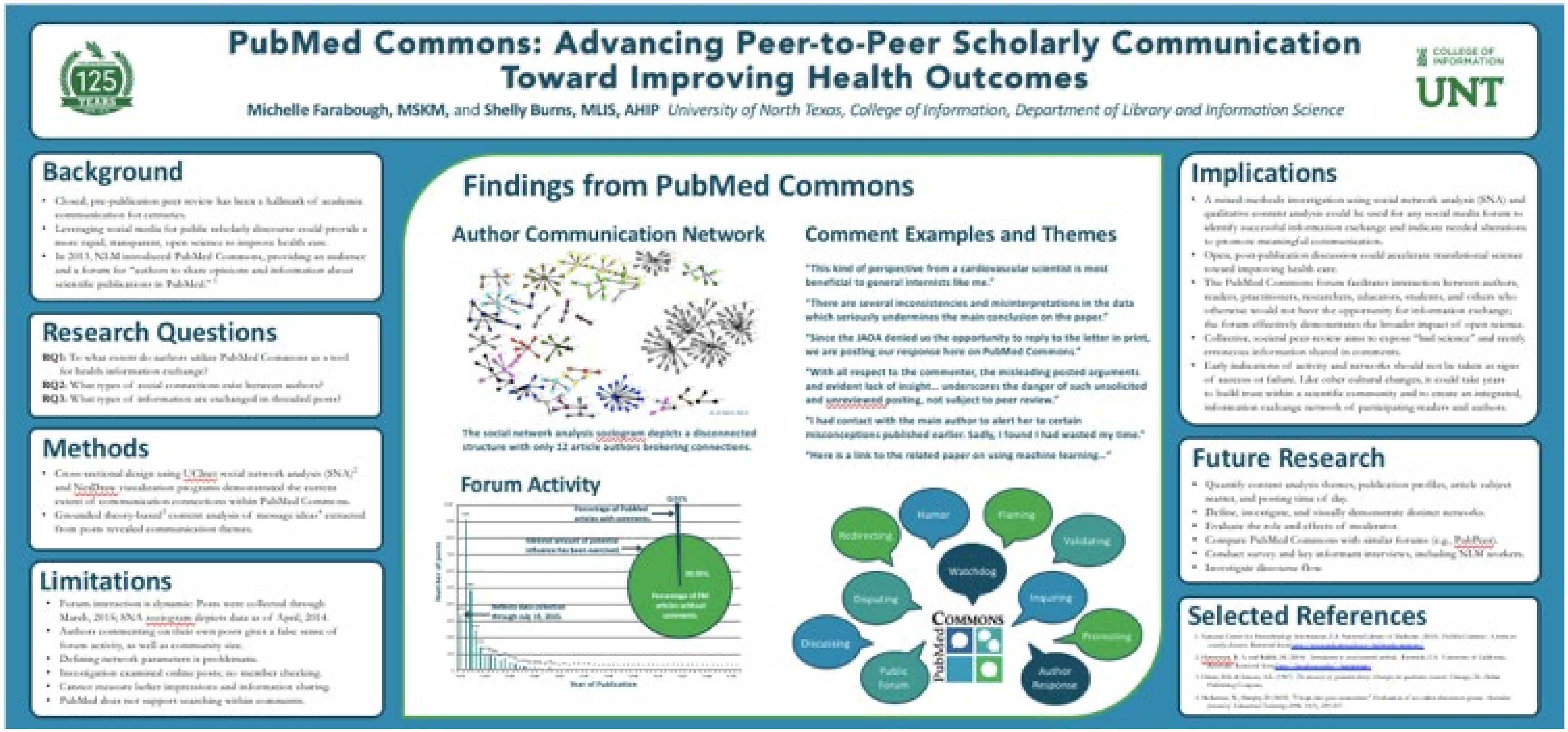
The Dedoose web application was utilized to create a database from a media file of all 7,629 PubMed Commons comments, 25 descriptor fields, and 11 qualitative content codes for comments.



APPENDIX D

POSTER PRESENTED AT THE CDC NATIONAL CONFERENCE

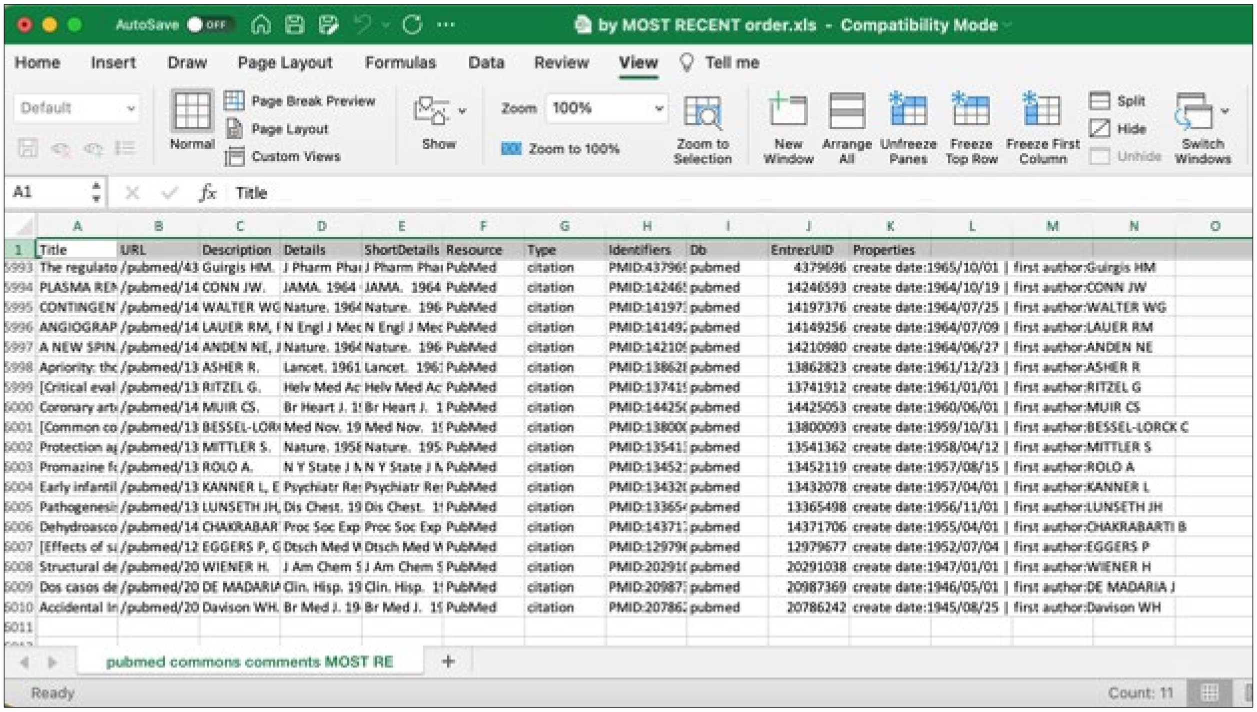
Farabough, M., & Burns, S. (2015). *PubMed Commons: Advancing peer-to-peer scholarly communication toward improving health outcomes.* Poster session at the 9th Annual CDC National Conference on Health Communication, Marketing and Media, Atlanta, GA.

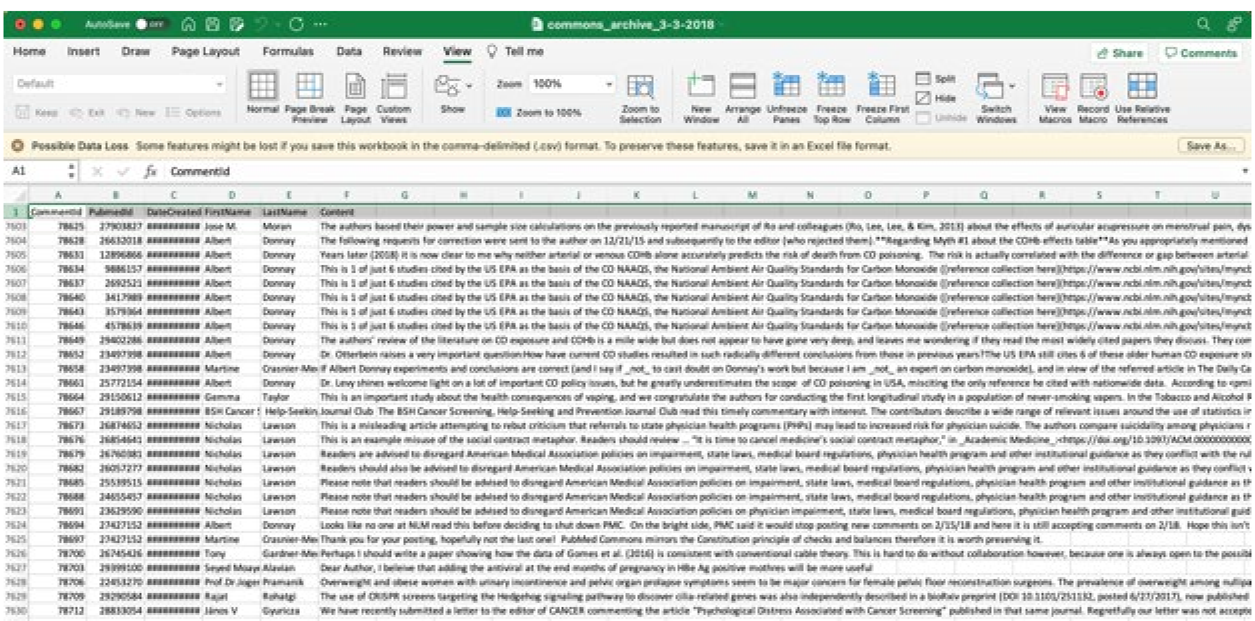


APPENDIX E

PUBMED COMMONS FORUM DATA

PubMed Commons data was collected from two comma separated value files: 1) PubMed search for “all[sb]” with “PubMed Commons Readers comments” filter activated (top) and 2) NCBI FTP website linked “commons_archive.csv” file (bottom).





APPENDIX F

KEY INFORMANT INTERVIEW QUESTIONS

Questions for meeting with PubMed Commons Editors Hilda Bastian and Melissa Vaught on October 28, 2016, at NLM offices—38A, Lister Hill Center, 10th floor, office 1003N in Bethesda, MD.

Questions:

1. Now that PubMed Commons has advanced from a pilot program, does management have explicit objectives for PubMed Commons beyond those espoused on the PubMed Commons website and in the PubMed Commons blog? (e.g., Build interdisciplinary ties? Promote professional profiles? Provide a platform for linking additional content? Ask questions of peers?)

2. Are there plans for additional features? (e.g., Allow scholars to build professional profiles? Private messaging, thus facilitating communication, but not public), Benefits to posters like Altmetric measures? Subscribe to threaded discussion? Social tagging? Upload video like Bush's Memex?)

3. Do you consider any other social networking sites competition? (e.g. PubPeer or forums on individual journals?)

4. Are you conducting quantitative research via some type of altmetrics/webometrics? If so, what software?

5. Are you conducting qualitative research on posts? If so, what type of methodology? (e.g., effects of anonymity?)

6. Are you collecting demographic data on posters? (e.g., geographic location? Tenure in an institution/organization? Discipline represented?

7. Are there plans for additional marketing? If so, what kind and where?

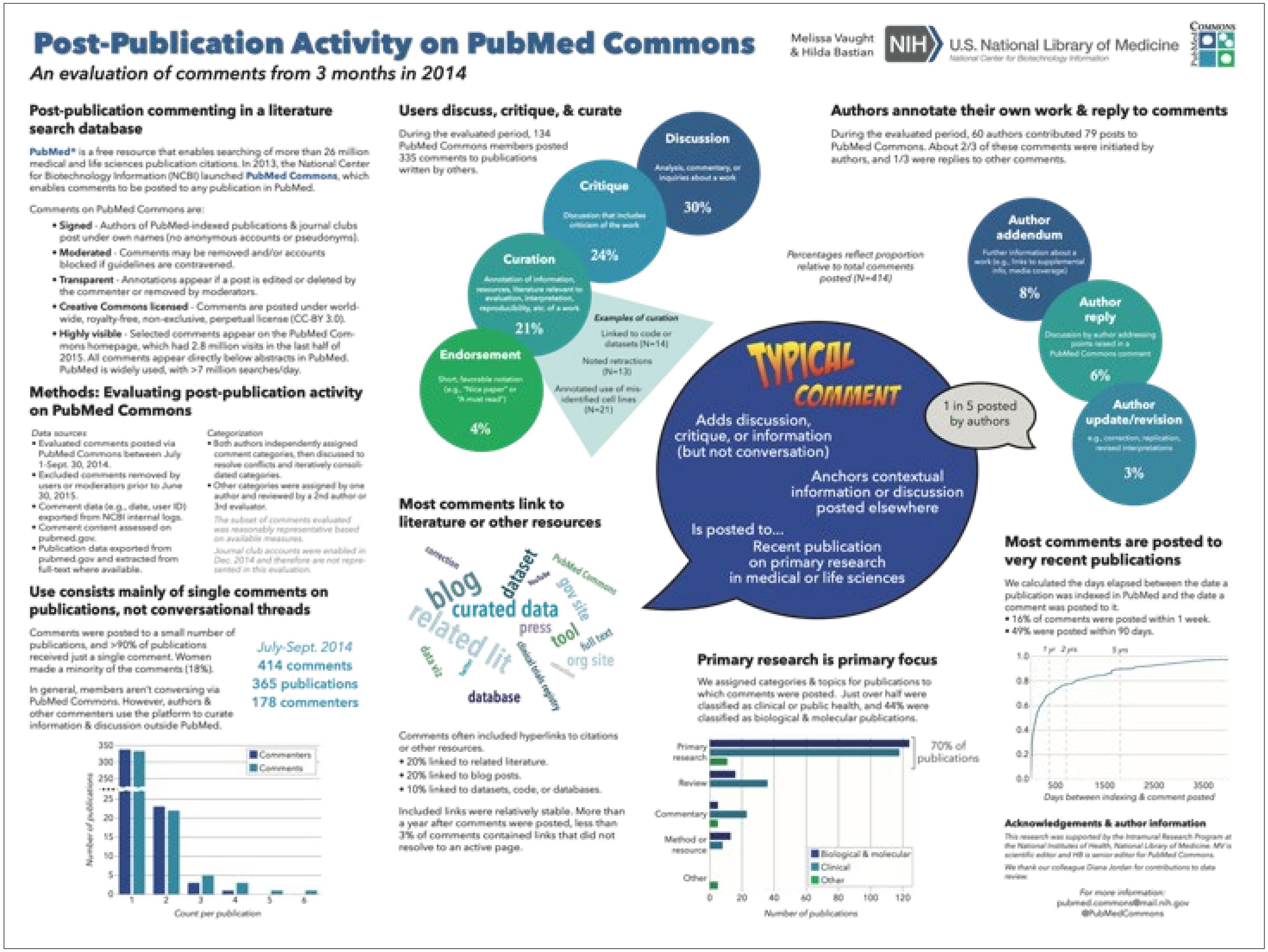
8. Do you have objections to me performing semi-structured interviews or questionnaires? If not, do you have a way for contacting them?

9. Do you have objections to me analyzing the data on the PubMed GUI that constitutes PubMed Commons?

APPENDIX G

POSTER PRESENTED AT THE AAAS ANNUAL MEETING

Vaught, M. & Bastina, H. (2016). *Post-publication Activity on PubMed Commons.* Poster session at the American Association for Advancement of Science (AAAS) Annual Meeting, Washington, D.C.



REFERENCES

Ali, P. & Watson, R. (2016). Peer review and the publication process. *Nursing Open,* 193-202. doi: 10.1002/nop2.51.

Allison, D. B., Brown, A. W., George, B. J., & Kaiser, K. A. (2016). Reproducibility: A tragedy of errors. *Nature, 530*(7588), 27-29. doi:10.1038/5300027a

Aluletta, K. (2009). *Googled: The end of the world as we know it.* Penguin Books.

Andrews, D., Nonnecke, B., & Preece, J. (2003). Electronic survey methodology: A case study in reaching hard to involve internet users. *International Journal of Human-Computer Interaction, 16*(2), 185-210.

Aragon, J. (2016). 502.3 Fundamental science practices: Peer Review. In USGS *Survey Manual*.https://www.usgs.gov/about/organization/science-support/survey-manual/5023-fundamental-science-practices-peer-review

Ascher, M. (2014, August 6). PubMed Commons: Set up alerts for comments on your papers in PubMed. *FYI The NYMC Health Sciences Library Blog.* http://fyi.libmedia.nymc.edu/?p=4549

Association for Information Science and Technology. (2021). *What is information science.* https://www.asist.org/about/what-is-information-science/

Association of College & Research Libraries. (2021). *Principles and strategies for the reform of scholarly communication.* http://www.ala.org/acrl/publications/whitepapers/principlesstrategies

Avwontom, F. (2016). "If you fail to plan, you are planning to fail." Can business plans reduce the high failure rate of SMEs? *eVALUation Matters, 2,* 64-73. http://idev.afdb.org/sites/default/files/Evaluations/2020-02/Article%20from_EMQ2__Fail%20to%20plan%20part%20one_WEB%207.pdf

Aziz, E. E. (2014). *How to secure 360 stakeholder buy-in and sustain it?* Paper presented at PMI® Global Congress 2014—North America, Phoenix, AZ. Newtown Square, PA: Project Management Institute. https://www.pmi.org/learning/library/secure-360-stakeholder-buy-in-9306

Ballantine, P. W., & Martin, B. A. (2005). Forming parasocial relationships in online communities. In G. Menon & A. R. Rao (Eds.) *NA-Advances in Consumer Research* (pp. 197-201). Association for Consumer Research.

Barab, S. A. (2003). An introduction to the special issue: Designing or virtual communities in the service of learning. *The Information Society, 19,* 197-201.

Barak, A. & Gluck-Ofri, O. (2007). Degree of reciprocity of self-disclosure in online forums. *CyberPsychology & Behavior, 10*(3). 407-417.

Barabasi, A-L. (2012). Luck or reason. *Nature, 489,* 507-508. https://doi.org/10.1038/nature11486

Bastian, H. (2013, October 22). Science buzz and criticism get a powerful boost. *Scientific American.* https://blogs.scientificamerican.com/absolutely-maybe/science-buzz-and-criticism-get-a-powerful-boost/

Bastian, H. (2014). A stronger post-publication culture is needed for better science. *PLoS Medicine, 11*(12), e1001772, 1-3. doi: 10.1371/journal.pmed.1001772

Bastian, H. (2017, March 15). *Evaluation of post-publication activities in PubMed* [webinar]. Webex Session 627725018.

Bastian, H. (2017). *Post-publication peer-review and certificate systems* [video]. Presentation at the National Academy of Sciences Arthur M. Sackler Colloquium on Reproducibility of Research: Issues and Proposed Remedies, Washington, D.C., March 8-10, 2017. *NCBI YouTube Channel.* https://www.youtube.com/watch?v=1sM8miLiYUY

Bauer, M. (2000). Classical content analysis. In M. Bauer & G. Gaskell (Eds.) *Qualitative researching with text, image and sound* (pp. 131-151). Sage Publications, Inc.

Baxter, R. L., De Reiemer, C., Landini, A., Leslie, L., & Singletary, M. W. (1985). A content analysis of music videos. *Journal of Broadcasting & Electronic Media, 29*(2), 333-340.

Baym, N. K. (1998). The emergence of on-line community. In S. Jones (Ed.) *Cybersociety: communication and community* (pp. 35-68). Sage Publications.

Beltrao, P. (2013). News. PubMed Commons: A system that enables researchers to share their opinions about scientific publications. *Biostars Bioinformatics Explained.* https://www.biostars.org/p/84222/

Bjork, B-C. & Solomon, D. (2012). Open access versus subscription journals: A comparison of scientific impact. *BMC Medicine, 10,* 73, doi: 10.1186/1741-7015-10-73

Blanchard, A., & Markus, M. L. (2002). Sense of virtual community – Maintaining the experience of belonging. In *Proceedings of the 35th annual Hawaii international conference on system sciences (HICSS)* (pp. 3566–3575). doi: http://dx.doi.org/10. 1109/HICSS.2002.994449.

Blooming Wellness. (2018, April 2). *Vitamin D: An interview with Dr. John Cannell, founder of the Vitamine D Council.* https://bloomingwellness.com/2018/04/should-you-supplement-with-vit-d-an-interview-with-dr-john-cannell-founder-of-the-vitamin-d-council/

BMJ Opinion. (2016, September 21). *William Cayley: What is your story?* https://blogs.bmj.com/bmj/2016/09/21/william-cayley-what-is-your-story/

Bordignon, F. (2020). Self-correction of science: A comprehensive study of negative citations and post-publication peer review. *Scientometrics, 124*, 1225-1239.

Borgatti, S.P., Mehra A., Brass, D.J., & Labianca, G., (2009). Network analysis in the social sciences. *Science, 323*(5916), 892-895. doi: 10.1126.science.1165821

Borgatti, S. P., Everett, M. G., & Johnson, J. C. (2012). *Analyzing social networks.* Sage Publications, Inc.

Borgman, C. L. (1990). Editor's introduction. In: Borgmann, C. L. ed. *Scholarly communication and bibliometrics.* (pp. 10-27). Sage Publications, Inc.

Borgman, C. L. (2000). Digital libraries and the continuum of scholarly communication. *Journal of Documentation, 56*(4), 412-430.

Bosman, J., & Kramer, B. (2015). 101 *Innovations in scholarly communication.* Poster presentation at Force2015 conference, Oxford, UK. https://figshare.com/articles/101_Innovations_in_Scholarly_Communication_the_Changing_Research_Workflow/1286826

Bravo, L. E. (2014). From printing to Scielo and PubMed Central. *Colombia Medica, 45*(1), 5-6.

Brown, J. S. & Duguid, P. (2000). *The social life of information.* Harvard Business School Publishing.

Brown, J. S. & Duguid, P. (1991). Organizational learning and communities of practice: Toward a unified view of working, learning, and innovation. *Organization Science, 2*(1), 40-57.

Bruckman, A. (2006). Teaching Students to Study Online Communities Ethically. *Journal of Information Ethics 15*(2), 82-98.

Bubber, M. (2013). *EVSC 200: Term paper research.* http://www.lib.sfu.ca/help/subject-guides/rem

Bubela, T., & Caulfield, T. (2009). Science communication reconsidered. *Nature Biotechnology, 27*(6), 514-518.

Burns, S. & Farabough, M. (2015). *PubMed Commons: Exploring PubMed's post-publication social media forum and its implications for the medical library community.* Paper presented at the Pacific Northwest Chapter of the Medical Library Association Annual Conference. Seattle, WA.

Burt, R. S. (1992). *Structural holes: The social structure of competition.* Harvard University Press.

Burt, R. S. (2000). *Structural holes versus network closure as social capital,* 1-30. https://snap.stanford.edu/class/cs224w-readings/burt00capital.pdf

Bush, V. (1945). As we may think. *The Atlantic, 7,* 112-124.

Bylieva, D., Lobatyuk, V., & Safonova, A. (2019). Online forums: Communication model, categories of online communication regulation and norms of behavior. *Humanities & Social Sciences Reviews, 7*(1), 332-340.

Cambridge University Press. (2021). Discussion. *Cambridge Dictionary.* https://dictionary.cambridge.org/us/dictionary/english/discussion

Cambridge University Press. (2021). Discourse. *Cambridge Dictionary.* https://dictionary.cambridge.org/us/dictionary/english/discourse

Carpenter, B. (n.d.) *About me.* https://bob-carpenter.github.io/

Casilli, A. A., Rouchier, J., & Tubaro, P. (2014). How to build consensus in a health-oriented online community: Modeling a "Pro-Ana" forum. *Revue francaise de sociologie (English Edition), 55*(4), 731-764.

Centers for Disease Control and Prevention. (2021). *CDC Social media tools, guidelines & best practices.* https://www.cdc.gov/socialmedia/tools/guidelines/index.html

Central Intelligence Agency. (1977). *Quantitative content analysis: A methodological handbook with examples from research on the Soviet Union.* https://www.cia.gov/library/readingroom/docs/DOC_0000498617.pdf

Cheung, W. S., Hew, K. F., & Ng, C. (2008). Toward an understanding of why students contribute in asynchronous online discussions. *Journal of Educational Computing Research, 38*(1), 29-50.

Christakis, N. A., & Fowler, J. H. (2009). *Connect4ed: How your friends; friends' friends affect everything you feel, think, and do*. Back Bay Books / Little, Brown and Company.

Chrome Web Store. (n.d.) *PubPeer.* https://chrome.google.com/webstore/detail/pubpeer/fmcdfigcdfkdghdklblbbpacikcchbbh?hl=en

CiCicco-Bloom, B., & Crabtree, B. F. (2006). The qualitative research interview. *Medical Education, 40,* 314-321.

Cobb, M. (2017). The prehistory of biology preprints: A forgotten experiment from the 1960s. *PLoS Biology, 15*(11): e2003995. https://doi.org/10.1371/journal.pbio.2003995

Cochrane. (2020). *You can make a difference.* https://crowd.cochrane.org/

Cockerill, M. (2013, November 29). Connecting up post-publication commentary via integration with PubMed Commons. *BMC Blog Network.* https://blogs.biomedcentral.com/bmcblog/2013/11/29/connecting-up-post-publication-commentary-via-integration-with-pubmed-commons/

Collins, F. S., & Tabak, L. A. (2014). NIH plans to enhance reproducibility. *Nature, 505*(7485), 612-613. https://www.nature.com/articles/505612a

Collins, F. (2014, August 5). PubMed Commons: Catalyzing scientist-to-scientist interactions. NIH Director's Blog. https://directorsblog.nih.gov/2014/08/05/pubmed-commons-catalyzing-scientist-to-scientist-interactions/

Constant, D., Sproull, L., & Kiesler, S. (1996). The kindness of strangers: The usefulness of electronic weak ties for technical advice. *Organization Science, 7*(2), 119-135.

Contractor N. S., & Eisenberg, E. M. (1990). Communication networks and new media in organizations. In J. Fulk & C. Steinfeld (Eds.) *Organizations and communication technology* (pp. 143-172). Sage Publications, Inc.

Cornell University. (n.d.) *arcXiv.org 30.* https://arxiv.org/

Cothrel, J., & Williams, R. (1999). On-line communities: Helping them form and grow. *Journal of Knowledge Management, 3*(1), 54-60.

Couchman, J. R. (2014). Peer review and reproducibility. Crisis or time for course correction? *Journal of Histochemistry & Cytochemistry, 62*(1), 9-10. doi: 10.1369/0022155413513462

Counterfactual, The. (2008, September 13). *About.* https://www.clivebates.com/about/

Craig, R. T., & Tracy, K. (1995). Grounded practical theory: The case of intellectual discussion. *Communication Theory, 5,* 248-272.

Crane, D. (1972). *Invisible colleges. Diffusion of knowledge in scientific communities.* The University of Chicago Press.

Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches.* Sage Publications, Inc.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches.* Sage Publications, Inc.

Creswell, J. W. (2014). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. (5th ed.). Pearson.

Dayton, A. (2006). Beyond open access: Open discourse, the next great equalizer. *Retrovirology, 3*(55). doi: 10.1186/1742-4690-3-55

de Solla Price, D. (1963). *Little science, big science.* Columbia University Press.

de Solla Price, D., & Beavers, D. (1966). Collaboration in an invisible college. *American Pyschologist, 9(*11), 1011-1018.

Dee, C. R. (2007). The development of the Medical Literature Analysis and Retrieval System (MEDLARS). *Journal of the Medical Library Association, 95*(4), 416-425.

Dictionary.com. (2021). *Sponsorship.* https://www.dictionary.com/browse/sponsorship

Digital Science. (n.d.) *The donut and Altmetric Attention Score.* https://www.altmetric.com/about-our-data/the-donut-and-score/

Dillon, J.T. (1990). *The Practice of Questioning.* Routledge.

Doren, K., & Jones, C. (2013). *That's Outside My Boat: Letting Go of What You Can't Control*. Andrews McMeel Publishing.

Driscoll, K. (2016). Social media's dial-up ancestor: The Bulletin Board System. *IEEE Spectrum.* Advance online publication https://spectrum.ieee.org/tech-history/cyberspace/social-medias-dialup-ancestor-the-bulletin-board-system

Duhigg, C. (2012). *The power of habit: Why we do what we do in life and business.* Random House.

Dynamic Ecology. (2021). *About.* https://dynamicecology.wordpress.com/about/

Eaton, A. (2016). *PubMed Commons API.* https://gist.github.com/hubgit/ed04da6ea8a2fca07583

École normale supérieure [ENS] Paris. (2021). *Boris Barbour's personal page.* https://www.ibens.ens.fr/spip.php?article296&lang=fr#pubpeer

Eisen, M. (2013, October 22). PubMed Commons: Post publication peer review goes mainstream. *it is NOT junk*. https://www.michaeleisen.org/blog/?p=1472

Elgin, S. H. (2004, August 13). *Discourse vs. discussion skills*. http://linguistlist.org/ask-ling/message-details1.cfm?asklingid=200326769

Europe PMC. (2018, December 4). *Bringing PubMed Commons to Europe PMC*. http://blog.europepmc.org/2018/12/bringing-pubmed-commons-to-europe-pmc.html

Eysenbach, G., Powell, J., Englesakis, M., Rizo, C., & Stern, A. (2004). Health related virtual communities and electronic support groups: A systematic review of the effects of online peer to peer interactions. *British Medical Journal, 328*(7449), 1166-1170. doi: http://dx.doi.org/10.1136/bmj.328.7449.1166

Falisi, A. L., Wiseman, K. P., Gaysynsky, A., Scheideler, J. K., Ramin, D. A., & Chou, W. S. (2017). Social media for breast cancer survivors: A literature review. *Journal of Cancer Survivors, 11,* 808-821. doi: 10.1007/s11764-017-0620-5.

Farabough, M., & Burns, S. (2015). *PubMed Commons: Advancing peer-to-peer scholarly communication toward improving health outcomes.* Poster session at the 9th Annual CDC National Conference on Health Communication, Marketing and Media, Atlanta, GA.

Feagin, J R., Orum, A. M. & Sjoberg, G. (1991). *A case for the case study.* University of North Carolina Press.

Fiorini, N., Lipman, D. J., & Zhiyong, L. (2017). Towards PubMed 2.0. *eLife, 6,* e28801. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5662282/

Fjällbrant, N. (1997). Scholarly communication: Historical development and new possibilities. *Proceedings of the IATUL Conferences.* http://docs.lib.purdue.edu/iatul/1997/papers/5

Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative Inquiry, 12,* 219-245. doi: 10.1177/1077800405284363

Foster Open Science. (n.d.) *What is open science? Introduction.* https://www.fosteropenscience.eu/content/what-open-science-introduction

Fronzetti Colladon, A., & Vagaggini, F. (2017). Robustness and stability of enterprise intranet social networks: The impact of moderators. *Information Processing and Management, 53*(6), 1287-1298.

Gabbriellini, S. (2014). The evolution of online forums as communication networks: An agent-based model, *Revue francaise de sociologie, 55*(4), 805-826.

Galbraith. D. W. (2015). Redrawing the frontiers in the age of post-publication review. *Frontiers in Genetics.* https://www.frontiersin.org/articles/10.3389/fgene.2015.00198/full

Gallagher, S. E., & Savage, T. (2013). Cross-cultural analysis in online community research: A literature review. *Computers in Human Behavior, 29*(3), 1028–1038. http://dx.doi.org/10.1016/j.chb.2012.09.011.

Garvey, W. D., & Griffith, B. C. (1972). Communication and information process within scientific disciplines: Empirical findings for psychology. *Information Storage and Retrieval, 8,* 123-126.

Gelman, A. (2013, October 23). PubMed Commons: A system for commenting on articles in PubMed. *Statistical Modeling, Causal Inference, and Social Science.*

https://statmodeling.stat.columbia.edu/2013/10/23/pubmed-commons-system-commenting-articles-pubmed/

Giles, D., Stommel, W., Paulus, T., Lester, J., & Reed, D. (2015). Microanalysis of online data: The methodological development of "digital CA." *Discourse, Context, and Media, 7,* 45-51.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research.* Aldine Publishing.

Glickman, S. W., Galhenage, S., McNair, L., Barber, A., Patel, K., Schulman, K. A., & McHutchison, J. G. (2012). The potential influence of internet-based social networking on the conduct of clinical research studies. *Journal of Empirical Research on Human Research Ethics: An International Journal, 7*(1), 71-80.

Goodman, S. (2014, April 22). *Professor Steven Goodman on PubMed Commons* [video]. https://www.youtube.com/watch?v=0j0I7tYTksw

Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology, 78*, 1360-1380.

Guba, E. G. & Lincoln, Y. S. (1994). Competing paradigms in qualitative research In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105–117). Sage Publications, Inc.

Hak, T., & Bernts, T. (1996). Coder training: Theoretical training or practical socialization? *Qualitative Sociology, 19*, 237-257.

Hamila, H. (n.d.) *Harri Hemila.* https://www.mv.helsinki.fi/home/hemila/education.htm

Hannemann, R. A. (2001). *Introduction to social network methods.* http://faculty.ucr.edu/~hanneman/nettext/

Hanneman, R. A. & Riddle, M. (2005a). Connections In *Introduction to social networks.* https://faculty.ucr.edu/~hanneman/nettext/C7_Connection.html

Hanneman, R. A. & Riddle, M. (2005b). Embedding In *Introduction to social networks.* https://faculty.ucr.edu/~hanneman/nettext/C8_Embedding.html

Hansen, D. L., Shneiderman, B., & Smith, M. A., (2011). *Analyzing social media networks with NodeXL: Insights from a connected world.* Elsevier.

Herring, S. (2004). Computer-mediated discourse analysis: An approach to researching online communities. In S.A. Barab, R. Kling, & J. H. Gray (Eds.) *Designing for virtual communities in the service of learning* (pp. 338-376). Cambridge University Press. doi: 10.1017/CB09780511805080.016.

Hasan, S. A., Masood, I., & Mumtaz, Z. (2016). Pubmed commons: Pros and cons. *Journal of the Pakistan Medical Association, 66*(7), 913. https://jpma.org.pk/article-details/7843

Haythornwaite, C. (1996). Social network analysis: An approach and set of techniques for the study of information exchange. *Library and Information Science Research, 18*(4), 323-342.

Haythornthwaite, C. (2002). Strong, weak, and latent ties and the impact of new media. *The Information Society, 18,* 385-401.

Hart, C. (1998). *Doing a literature review: Releasing the social science research imagination.* Sage Publications, Inc.

Hasan, S. A., Masood, I., & Mumtaz, Z. (2016). PubMed Commons: Pros and cons. *Journal of Pakistan Medical Association, 66*(7), 913. http://www.jpma.org.pk/full_article_text.php?article_id=7843

Heckman, R., & Annabi, H. (2006). How the teacher's role changes in on-line case study discussions. *Journal of Information Systems, 17*(2), 141-150.

Heneghan, C. (2013, October 23). What you ought to know about PubMed Commons. *Carl Heneghan… Finding the Truth in Research that Matters.* http://www.carlheneghan.com/what-you-ought-to-know-about-pubmed-commons/

Herron, D. M. (2011). Is expert peer review obsolete? A model suggests that post-publication reader review may exceed the accuracy of traditional peer review. *Surgical Endoscopy, 26,* 2275-2280.

Hew, K.F. & Cheung, W.S. (2003). Evaluating the participation and quality of thinking of pre-service teachers in an asynchronous online discussion environment: Part I. *International Journal of Instructional Media, 30*(3), 247-263.

Hildebrand, M., Ahumnada, C., & Watson, S. (2013). CrowdOutAIDS: Crowdsourcing youth perspectives for action. *Reproductive Health Matters, 21*(41), 57-68.

Hilty, L., & Hercheui, M. D., (2010). ICT and sustainable development. In J. Berleur, M. D. Hercheui, & L. Hilty (Eds.) *What kind of information society? Governance, virtuality, surveillance, sustainability, resilience. Proceedings of the 9th Human Choice* and *Computers International Conference (HCC9) and the WCC 2010 Critical Infrastructure Protections Conference*. Springer.

Hine, C. (2000). Virtual Ethnography. Sage Publications, Inc.

Hurd, J. (2000). The transformation of scientific communication: A model for 2020. *Journal of the American Society for Information Science, 51*(14), 1279-1283.

Institute of Medicine. (2009). The value, importance, and oversight of health research. In Nass S. J., Levit, L. A., & Gostin L. O. (Eds.). *Beyond the HIPAA privacy rule: Enhancing privacy, improving health through research* (pp. 111-152). National Academies Press. doi: https://doi.org/10.17226/12458

InterTAC Information Specialists' Sub-Group [ISSG]. (2021). *The InterTAC Information Specialists' Sub-Group Search Filter Resource.* https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home?authuser=0

Jonassen, D., & Remidez, J. (2005). Mapping alternate discourse structures onto computer conferences. International *Journal of Knowledge and Learning, 1*(1/2), 113-129.

Journal of the Canadian Dental Association. (2013, December 4). *PubMed pilots open peer review system*. https://jcda.ca/article/d194

Kahn, S. (2009). Wonderful wikis and internet forums. *Science and Children, 46*(9), 27-31.

Kaiser, C., & Bodendorf, F. (2012). Mining consumer dialog in online forums. *Internet Research, 22*(3), 275-297.

Kaminski, J. (2011). Diffusion of Innovation Theory. *Canadian Journal of Nursing Informatics, 6*(2), 1-7. https://cjni.net/journal/?p=1444

Kao, S. C., Wu, C., & Huang, C. L. (2017). KCE_{BSA}: Development of online knowledge community evaluation model using the balanced scorecard approach. *Journal of Enterprise Information Management, 30*(4), 625-643.

Katz, M. H., & Redberg, R. F. (2018). Science requires open discourse. *JAMA Internal Medicine, 178*(1), 15-16.

Kelly, J., Sadeghieh, T., & Adeli, K. (2014). Peer review in scientific publications: Benefits, critiques, and a survival guide. *The Journal of the International Federations of Clinical Chemistry and Laboratory Medicine, 25*(3), 227-243.

Kimmerle, J., Thiel, A., Gerbing, K., Bientzle, M., Halatchliyski, I., & Cress, U. (2013). Knowledge construction in an outsider community: Extending the communities of practice concept. *Computers in Human Behavior, 29,* 1078-1090.

Kinnunen, J. (1996). Gabriel Tarde as a founding father of innovations diffusion research. *Acta Sociologica, 39*(4), 431-442.

Kline, P., & Saunders, B. (1993). *Ten steps to a learning organization.* Great Ocean Publishers, Inc.

Koku, D. F., & Wellman, B. (2002). Scholarly networks as learning communities: The case of TechNet. In S. Barab, R. Kling, & J. Gray (Eds.) *Designing virtual communities in the* *service of learning.* (Learning in doing: Social, cognitive and computational perspectives, pp. 299-337). Cambridge University Press. doi:10.1017/CBO9780511805080.015 Cambridge University Press.

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology.* Sage Publications, Inc.

Knoepfler, P. (2015). Reviewing post-publication peer review. *Trends Genet, 31*(5), 221-223.

Kuhn, T. S. (1962). *The structure of scientific revolutions.* University of Chicago Press.

Lampe, C., Zube, P., Lee., J., Park., C. H., & Johnston, E. (2014). Crowdsourcing civility: A natural experiment examining the effects of distributed moderation in online forums. *Government Information Quarterly, 31*(2), 317-326.

Landhuis, E. (2016, July 20). Scientific literature: Information overload. *Nature, 535*, 457-458. https://www.nature.com/articles/nj7612-457a

Lane, P. (2016a). *Adoption of PubMed Commons as a forum for post-publication peer review* [Poster presentation]. International Society for Medical Publication Professionals European Meeting, January 20-21, London, England, United Kingdom. http://www.ismpp.org/assets/docs/Education/AnnualMeeting/11thAM/Posters/26_lane_ismpp11am_poster.pdf

Lane, P. (2016b). *Adoption of PubMed Commons as a forum for post-publication peer review* [Poster presentation]. International Society for Medical Publication Professionals Annual Meeting, , April 27-29. Arlington, VA, United States.

Larson, R. & Larson, E. (2009). Top five causes of scope creep... and what to do about them. Paper presented at PMI® Global Congress 2009—North America, Orlando, FL. Newtown Square, PA. https://www.pmi.org/learning/library/top-five-causes-scope-creep-6675

Lave, J. & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge University Press.

Li, R., Huang, M. & Zhu, X. (2010). Sentiment analysis with global topics and local dependency. In *Proceedings of AAAI* (pp. 1371-1376).

Lindberg, D. A. B. (2011). The National Library of Medicine. *World Neurosurgery, 74*(1), 46-48. doi: 10.1016/j.wneu.2010.04.014

Lindenbaum, P. (2014). New: PubMed Commons: A system that enables researchers to share their opinion about scientific publications. *Biostars Bioinformatics Explained.* https://www.biostars.org/p/84222/

Ling, K., Beenen, G., Ludford, P.J., Wang, X., Chang, K., Li, X., … & Kraut, R. (2005). Using social psychology to motivate contributions to online communities. *Journal of Computer Mediated Communication, 10*(4). 10.1111/j.1083-6101.2005.tb00273.x

Lizarondo, L., Kumar, S., & Grimmer-Somers, K. (2010). Online journal clubs: An innovative approach to achieving evidence-based practice. *Journal of Allied Health, 39*(1), e17-e22.

Lukas, J. A. (2014). Project management: Getting back to basics. Paper presented at PMI® Global Congress 2014—North America, Phoenix, AZ. Newtown Square, PA. https://www.pmi.org/learning/library/project-management-getting-back-basics-9313

Madden, A. D. (2000). When did peer review become anonymous? *Aslib Proceedings, 52*(8), 273-276.

Malinen, S. (2015). Understanding user participation in online communities: A systematic literature reviw of empirical studies. *Computers in Human Behavior, 46,* 228-238.

Manca, S., Delfino, M., & Mazzoni, E. (2009). Coding procedures to analyse interaction patterns in educational web forums. *Journal of Computer Assisted Living, 25*(2). doi: 10.1111/j.1365-2729.2009.00296.x

Marion, L. S., Garfield, E., Hargens, L. L., Lievrous, L. A., White, H. D., & Wilson, C. S. (2003). Social network analysis and citation analysis: Complimentary approaches to the study of scientific communication. *Proceedings of the American Society for Information Science and Technology, 40*(1), 496-497.

McCracken, D. (1988). *The long interview* (Qualitative Research Methods, Vol. 13). Sage Publishing, Inc.

Meadows, A. J. (1974). *Communication in science.* Butterworths.

Meadows, A. J. (1998). *Communicating research.* Academic Press.

Merriam-Webster. (n.d.). *peer.* https://www.merriam-webster.com/dictionary/peer

METRICS. (n.d.). *Robert Tibshirani*. https://metrics.stanford.edu/people/robert-tibshirani

McKenzie, R. D (2017). *Closing online comments: A case study of News24* (ISSN-L 2310-7855). [master's thesis. Stellenbosch University]. SUNScholar. http://scholar.sun.ac.za/handle/10019.1/101165.

Milano, B. (2019, February 7). Big tech's power growing at runaway speed. *The Harvard Gazette.* https://news.harvard.edu/gazette/story/2019/02/government-cant-keep-up-with-technologys-growth/

Milgram, S. (1967). The small world problem. *Psychology Today, 1,* 60-67.

Miller, K. I. (2000). Common ground from the post-positivist perspective: From “straw person” argument to collaborative coexistence. In S. R. Corman & M. S. Poole’s *Perspectives on organizational communication: Finding common ground*, (pp. 46-57). The Guildford Press.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Journal of Clinical Epidemiology, 62*(10), 1006-1012. doi:10.1016/j.jclinepi.2009.06.005

Mokoena, S. (2013). Engagement with and participation in online discussion forums. *The Turkish Online Journal of Education Technology, 12*(2). 97-105.

Montes, J., & Butler, D. L. (2008). Debating race through the tourist plantation: Analyzing a *New York Times* conversation. *Southeaster Geographer, 48*(3), 303-315.

Morzy, M. (2013). Evolution of Online Forum Communities. In T. Ozyer, J. Rokne, G. Wagner, & A H. P. Reuser’s (Eds.) *The Influence of technology on social network analysis and mining* (pp. 623-628). Springer.

Moscow State University (n.d.) *Sergei Ostroumov.* https://moscowstate.academia.edu/SergeiOstroumov

Muller, D. (2019). Open discourse: A media theory for the twenty-first century. *Ethical Space: The International Journal of Communication Ethics, 16*(1), 3-10.

Mumford, L. (1934). *Technics and civilization*. Harcourt, Brace and Co.

Mumford, L. (1959). An appraisal of Lewis Mumford’s ‘Technics and Civilization’ (1934). *Daedalus, 88,* 529.

Nassi-Calo, L. (2013, December 20). PubMed Commons: NLM launches pilot version of open comments on articles. *SciELO in Perspective.* https://blog.scielo.org/en/2013/12/20/pubmed-commons-nlm-launches-pilot-version-of-open-comments-on-articles/#.YSqfHNNKiLq

National Academies of Sciences, Engineering, and Medicine. (2017). *Strengthening the disaster resilience of the academic biomedical research community: Protecting the nation’s investment*. Washington, DC: National Academies Press. https://www.ncbi.nlm.nih.gov/books/NBK464163/

National Institutes of Health (NIH). (2020a). *National Library of Medicine (NLM).* https://www.nih.gov/about-nih/what-we-do/nih-almanac/national-library-medicine-nlm

National Institutes of Health (NIH). (2020b). *What we do.* https://www.nih.gov/about-nih/what-we-do/budget

Naughton, J. (2012). *From Gutenberg to Zuckerberg.* Quercus.

NCBI Insights. (2013, October 22). *PubMed Commons: A new forum for scientific discourse*. https://ncbiinsights.ncbi.nlm.nih.gov/2013/10/22/pubmed-commons-a-new-forum-for-scientific-discourse/

NCBI Insights. (2013, October 23). *Joining PubMed Commons: A step-by-step guide.* https://ncbiinsights.ncbi.nlm.nih.gov/2013/10/23/joining-pubmed-commons-a-step-by-step-guide/#more-315

NCBI Insights. (2013, November 1). *Early developments in the PubMed Commons pilot.* https://ncbiinsights.ncbi.nlm.nih.gov/2013/11/01/early-developments-in-the-pubmed-commons-pilot/

NCBI Insights. (2017, June 30). *July 12th NCBI Minute "Crowdsourcing post-publication comments: How your journal club can contribute using PubMed Commons"*. https://ncbiinsights.ncbi.nlm.nih.gov/2017/06/30/july-12th-ncbi-minute-crowdsourcing-post-publication-comments-how-you-and-your-journal-club-can-contribute-using-pubmed-commons/

NCBI Insights. (2018, February 27). *PubMed Commons to be discontinued.* Retrieved from https://ncbiinsights.ncbi.nlm.nih.gov/2018/02/01/pubmed-commons-to-be-discontinued/

NCBI Resource Coordinators. (2014). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research, 43*(D1), D6-D17. https://academic.oup.com/nar/article/43/D1/D6/2438293?login=true

National Center for Biotechnology Information (NCBI). (n.d.a). PubMed overview. https://pubmed.ncbi.nlm.nih.gov/about/

National Center for Biotechnology Information (NCBI). (n.d.b). *Our mission.* https://www.ncbi.nlm.nih.gov/home/about/mission/

National Center for Biotechnology Information (NCBI). (2013, November 14). https://pubmedcommonsblog.ncbi.nlm.nih.gov/ [no longer available]

National Center for Biotechnology Information (NCBI). (2013, November 26). PubMed Commons going public soon. https://pubmedcommonsblog.ncbi.nlm.nih.gov/ [no longer available]

National Center for Biotechnology Information (NCBI). (2018, February 1). PubMed Commons to be discontinued. https://ncbiinsights.ncbi.nlm.nih.gov/2018/02/01/pubmed-commons-to-be-discontinued/

National Institutes of Health. (2017). National Library of Medicine (NLM). *The NIH Almanac.* https://www.nih.gov/about-nih/what-we-do/nih-almanac/national-library-medicine-nlm

National Institutes of Health Office of Extramural Research. (2017). *NIH grants policy statement.* https://grants.nih.gov/grants/policy/nihgps/HTML5/section_8/8.2_availability_of_research_results_publications__intellectual_property_rights__and_sharing_research_resources.htm

National Library of Medicine (NLM). (n.d.). *PubMed Overview.* https://pubmed.ncbi.nlm.nih.gov/about/

National Library of Medicine (NLM). (2015). *Yearly citation totals from 2013 MEDLINE/PubMed baselines: 21,508, 439 citations found.* https://www.nature.com/articles/nj7612-457a

National Library of Medicine (NLM). (2016). *FAQ: PubMed.* https://www.nlm.nih.gov/services/pubmed_searches.html

National Library of Medicine (NLM). (2017a). *Fact sheet: PubMed: MEDLINE retrieval on the world wide web.* https://www.nlm.nih.gov/pubs/factsheets/pubmed.html

National Library of Medicine (NLM). (2017b). *Fact sheet: MEDLINE, PubMed, and PMC (PubMed Central): How are they different?* https://www.nlm.nih.gov/pubs/factsheets/dif_med_pub.html

National Library of Medicine (NLM). (2017c). *Fact sheet: MEDLINE.* https://www.nlm.nih.gov/pubs/factsheets/medline.html

National Library of Medicine (NLM). (2018). *PubMed Commons.* https://www.ncbi.nlm.nih.gov/pubmedcommons/ [No longer available]

National Library of Medicine (NLM). (2019). *Yearly citation totals from 2018 MEDLINE/PubMed baselines: 27,837,540 citations found.* https://www.nlm.nih.gov/bsd/licensee/2018_stats/2018_Totals.html

National Library of Medicine (NLM). (2020). *Citations added to MEDLINE by fiscal year.* https://www.nlm.nih.gov/bsd/stats/cit_added.html

National Library of Medicine (NLM). (2021). *Medline overview.* https://www.nlm.nih.gov/medline/medline_overview.html

National Science Foundation. (n.d.) *Public access: The results of NSG-funded research.* https://www.nsf.gov/news/special_reports/public_access/about_repository.jsp

Neuendorf, K. A. (2002). *The content analysis guidebook*. Sage Publishing, Inc.

Niewiadomski, P. (2013, October 29). *Why PubMed Commons will fail*. http://www.pawelthebiologist.com/pubmed-commons-will-fail/

NLM In Focus. (2016, June 30). *PubMed Celebrates its 20th anniversary!* https://infocus.nlm.nih.gov/2016/06/30/pubmed-celebrates-its-20th-anniversary/

NLM Technical Bulletin. (2017). *NNLM webinar: "What's new in post-publication activities in PubMed" on March 15, 2017.* https://www.nlm.nih.gov/pubs/techbull/ma17/brief/ma17_nnlm_webinar_post-publication.html

OpenTrials. (n.d.) *About.* https://opentrials.net/index.html

Oransky, I. (n.d.) *Meet PubPeer 2.0: New version of post-publication peer review site launches today.* https://retractionwatch.com/2017/06/15/meet-pubpeer-2-0-new-version-post-publication-peer-review-site-launches-today/

Oransky, I. (2013, October 22). PubMed now allows comments on abstract—but only a select few. *Retraction Watch.* https://retractionwatch.com/2013/10/22/pubmed-now-allows-comments-on-abstracts-but-only-by-a-select-few/

Oransky, I. (2017, June 15). Meet PubPeer 2.0: New version of post-publication peer review site launches today . *Retraction Watch.* https://retractionwatch.com/2017/06/15/meet-pubpeer-2-0-new-version-post-publication-peer-review-site-launches-today/

Orr, J. (1996). *Thinking about machines*. Cornell University Press.

Oslo University Hospital. (n.d.). *Morten Pedersen Oksvold.* https://ous-research.no/home/peng/Group+members/6038

Preece, J. & Maloney-Krichmar, D. (2005). Online communities: Design, theory, and practice. *Journal of Computer-Mediated Communication, 10*(4), doi: 10.1111/j.1083-6101.2005.tb00264.x

PLoS Medicine Editors. (2005). Meeting the needs of different research communities. *PLoS Medicine, 2*(5), e158.

Prendergast, G., Ko., D., & Yin, V. (2010). Online word of mouth and consumer purchasing intentions. *International Journal of Advertising. 29*(5), 687-708.

Procter, R., Williams, R., Stewart, J., Poschen, M., Snee, H., Voss, A., & Asgari-Targhi, M. (2010). Adoption and use of Web 2.0 in scholarly communications. *Philosophical Transactions of the Royal Society, 368*(1926), 4039-4056. Doi: 1098/rsta.2010.0155.

PubMed Commons. (2013). FAQ. PubMed Commons Blog. http://www.ncbi.nlm.nih.gov/pubmedcommons/faq/#joincommons [no longer available]

PubMed Commons Team. (2014, December 17). Introducing PubMed Commons Journal Clubs. PubMed Commons Blog. https://pubmedcommonsblog.ncbi.nlm.nih.gov/2014/12/17/introducing-pubmed-commons-journal-clubs/ [no longer available]

PubMed Commons Team. (Jul 5, 2017). Collaborating to bring journal clubs to PubMed Commons: A librarian's perspective. PubMed Commons Blog. [no longer available]

PubMed Commons Team. (2016, November 21). PubMed comments and their continuing conversations. PubMed Commons Blog. https://pubmedcommonsblog.ncbi.nlm.nih.gov/2016/11/21/pubmed-comments-their-continuing-conversations/ [no longer available]

PubMed Commons Team. (2017a, April 6). Author alerting readers via PubMed Commons. PubMed Commons Blog. https://pubmedcommonsblog.ncbi.nlm.nih.gov/2017/04/06/authors-alerting-readers-via-pubmed-commons/ [no longer available]

PubMed Commons Team. (May 28, 2017). Critiquing systematic review search strategies on PubMed. PubMed Commons Blog. [no longer available]

PubMed Commons Team. (2017b, July 5). Collaborating to bring journal clubs to PubMed Commons: A librarian's perspective. PubMed Commons Blog. https://pubmedcommonsblog.ncbi.nlm.nih.gov/ [no longer available]

PubPeer Foundation. (2021). *About PubPeer.* https://pubpeer.com/static/about

Rainie, L., Anderson, J.., & Albright, J. (2017, March 29). Theme 4: Oversight and community moderation come with a cost. The Future of free speech, trolls, anonymity and fake news online. *Pew Research Center.* https://www.pewresearch.org/internet/2017/03/29/theme-4-oversight-and-community-moderation-come-with-a-cost/

Ramos, J., Davies, G., Grant, T., McMinn, D., Nunn, J., & Wilson, E. (2015). *An evaluation of the utility and impact of PubMed Commons during the pilot phase.* Poster session presented at the International Society for Medical Publication Professionals, Arlington, VA.

Ramos, J., Davies, G., Grant, T., McMinn, D., Nunn, J., & Wilson, E. (2016, January 20-21). *An evaluation of the utility and impact of PubMed Commons during the pilot phase.* International Society for Medical Publication Professionals European Meeting, London, England, United Kingdom.

Ransom, D. C., La Gurdia, J. G., Woody, E. Z., & Boyd, J. L. (2010). Interpersonal interactions on online forums addressing eating concerns. *International Journal of Eating Disorders, 43*(2), 161-170.

Regenberg, A. (2010). Tweeting science and ethics: Social media as a tool for constructive public engagement. *American Journal of Bioethics, 10*(5), 30-31.

Research Gate. (2021). *Amanda Capes-Davis.* https://www.researchgate.net/profile/Amanda-Capes-Davis

Retraction Watch. (n.d.) *About Ivan Oransky.* https://retractionwatch.com/meet-the-retraction-watch-staff/about/

Riegelman, A., & Bakker, C. (2018). Understanding the complexities of retractions: Recommended resources. *College & Research Libraries News, 79*(1). 38-42.

Rigby, K., Sutherland, J., & Takeuchi, H. (2016). Embracing agile: How to master the process that's transforming management. *Harvard Business Review, 94*(5), *40-48, 50.* https://hbr.org/2016/05/embracing-agile

Rock, C. (Dec. 18. 2018). How to stay in control in the ever-changing world of technology. *Forbes.* https://www.forbes.com/sites/forbestechcouncil/2018/12/18/how-to-stay-in-control-in-the-ever-changing-world-of-technology/?sh=6964e2f32789

Rogers, E. (1995). *Diffusion of innovations* (4th Ed.). Free Press.

Rogers, E. M. & Kincaid, D.L. (1981). *Communication networks: Toward a new paradigm for research.* Free Press.

Romiszowski, A., & Mason, R. (2004). Computer-mediated communication. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (pp. 397-432). Lawrence Erlbaum.

Rong, L. Q., Lopes, A. J., Hameed, I., Gaudino, M., & Charlson, M. E. (2020). Examining the correlation between Altmetric score and citation count in the anaesthesiology literature. *British Journal of Anaesthesia, 125*(2), e223-e226.

Rubin, H. J. & Rubin, I.S. (2012). *Qualitative interviewing: The art of hearing data.* Sage Publishing, Inc.

Saravanakumar, M., & SuganthaLakshmi, T. (2012). Social media marketing. *Life Science Journal, 9*(4), 4444-4451.

Saunders, N. (2016). *PubMed Commons.rb.* https://github.com/neilfws/PubMed/blob/master/commons/code/ruby/pubmedcommons.rb

Saunders, N. (2016, December 1). An analysis of contributions to PubMed Commons. *R-bloggers.* https://www.r-bloggers.com/2016/12/an-analysis-of-contributions-to-pubmed-commons/

Saunders, N. (2018, March 14). *Analysis of contributions to PubMed Commons.* https://github.com/neilfws/PubMed/blob/master/commons/code/R/pubmedcommons.md

Saunders, N. (2018, March 15). Farewell then, PubMed Commons. *End to End Scalable Data Science with R, Rstudio R Markdown, Shiny, Keras, TensorFlow, sparklyr…* https://www.r-craft.org/r-news/farewell-then-pubmed-commons/

Schlager, M. S., Fusco, J., & Schank, P. (2002). Evolution of an online education community of practice. In K.A. Renniger & W. Shumar (Eds.), *Building virtual communities: Learning and change in cyberspace* (129-158). Cambridge University Press. https://books.google.com/books?hl=en&lr=&id=GY195ZNrLZ0C&oi=fnd&pg=PA129&dq=history+of+online+forums+for+education&ots=UKl4a-5brt&sig=28J7fH8FJelfm1QQxRjLuMAXMxY#v=onepage&q&f=false

Schultz-Jones, B. (2009). Examining information behavior through social networks: An interdisciplinary review. *Journal of Documentation, 65*(4), 592-631.

Science News Staff. (2016, December 7). PubPeer winds closed watched legal battle over anonymous comments. *Science.* http://www.sciencemag.org/news/2016/12/pubpeer-wins-closely-watched-legal-battle-over-anonymous-comments

Seo, K. K. (2006). Coming out of the darkness: Utilizing a systematic approach to evaluate online discussion forums in light of learning objectives. *Educational Technology, 46*(6), 24-28.

Shiode, J. H., & Parriott, J. R. (2016). Astronomy and Astrophysics. In *AAAS report XL: Research and development FY 2016.* https://www.aaas.org/fy16budget/astronomy-and-astrophysics

SimilarWeb. (2019, August). Organic Search Report. info@similarweb.com

SimilarWeb. (2019, August). Referral Traffic Report. info@similarweb.com

SimilarWeb. (2019, August). Strategic Overview.

Slomski, A. (2011). The National Library of Medicine: 175 years of advancing biomedical knowledge. *Journal of the American Medical Association, 305*(21), 2158-2161.

Smith K. (2013). A Brief History of NCBI's Formation and Growth. In: *The NCBI Handbook* (2nd ed.). National Center for Biotechnology Information. https://www.ncbi.nlm.nih.gov/books/NBK148949/

Spector, R. (2013, October 29). Stanford professor's pivotal role in bringing commenting capability to PubMed. *Stanford Medicine News Center.* https://med.stanford.edu/news/all-news/2013/10/stanford-professors-pivotal-role-in-bringing-commenting-capability-to-pubmed.html

Spencer, R., Pryce, J., and Walsh, J. (2014). Philosophical Approaches to Qualitative Research. In *The Oxford Handbook of Qualitative Research Methods* (pp. 81-98). http://dx.doi.org/10.1093/oxfordhb/ 9780199811755.001.0001

Stanford GSB Staff. (2011, February 1). When is bad publicity good? *Stanford Graduate School of Business.* gsb.stanford.edu/insights/when-bad-publicity-good

Stanoevska-Slabeva, K., Sacco, V., & Giardina, M. (2012). Content curation: A new form of gatewatching for social media? *Academia,* 1-35. https://bit.ly/3yaXcox

Steuber, K. R., & Solomon, D. H. (2008). Relational uncertainty, partner interference, and infertility: A qualitative study of discourse within online forums. *Journal of Social and Personal Relationships, 25*(5), 831-855.

Sun, A., Lin, C., Wu, M., Zhou, J., & Luo, L. (2018). A tale of two communication tools: Discussion-forum and mobile instant-messaging apps in collaborative learning. *British Journal of Education Technology, 49*(2), 248-261.

Swartz, A. (2013).Post-publication peer review mainstreamed. *Daily News' The Scientist.* https://www.the-scientist.com/daily-news/post-publication-peer-review-mainstreamed-38529

Swoger, B. (2014). Post-publication peer-review: Everything changes, and everything stays the same. *Scientific American.* https://blogs.scientificamerican.com/information-culture/post-publication-peer-review-everything-changes-and-everything-stays-the-same/

Szulanski, G. (2000). The process of knowledge transfer: A diachronic analysis of stickiness. *Organizational Behavior and Human Decision Processes, 82*(1), 9-27.

Tannery, N. (2014). PubMed Commons: A forum for PubMed authors. *University of Pittsburgh HSLS Update.* https://info.hsls.pitt.edu/updatereport/2014/february-2014/pubmed-commons-a-forum-for-pubmed-authors/

Teixeira da Silva, J. A., Al-Khatib, A., & Dobranszki, J. (2017). Fortifying the corrective nature of post-publication peer review: Identifying weaknesses, use of journal clubs, and rewarding conscientious behavior. *Science and Engineering Ethics, 23,* 1213-1226. doi: 10.1007/s11948-016-9854-2

Tennant, J. (2017, April 12). What are the barriers to post-publication peer review? https://blogs.lse.ac.uk/impactofsocialsciences/2017/04/12/what-are-the-barriers-to-post-publication-peer-review/

Tibshirani, R. (2013). PubMed Commons, A system for commenting on articles in PubMed. http://statweb.stanford.edu/~tibs/PubMedCommons.html

Tirado-Morueta, R., Maraver-Lopez, O., & Hernando-Gomez, A. (2017). Patterns of participation and social connections in online discussion forums. *Small Group Research, 48*(6), 639-664.

Udell, J. (2018, February 9). Preserving comments from PubMed Commons. *Hypothes.is.* https://web.hypothes.is/blog/archiving-pmc-comments/

UNESCO Institute for Statistics. (n.d.). *Information and communication technologies (ICT).* http://uis.unesco.org/en/glossary-term/information-and-communication-technologies-ict

University of California Berkeley. (n.d.) *Jennifer A. Doudna.* https://vcresearch.berkeley.edu/faculty/jennifer-doudna

University of California Museum of Paleontology, Berkeley. (2018). The scientific community: Diversity makes the difference. In *Understanding science.* https://undsci.berkeley.edu/article/socialsideofscience_02

University of Oxford Medical Sciences Division. (2021). *Bed Goldacre.* https://www.phc.ox.ac.uk/team/ben-goldacre

University of Washington Libraries. (n.d.). *Publication cycle* [digital image]. https://www.lib.washington.edu/scholpub/images/publicationcyclev2.png/image_view_fullscreen

University of Winnipeg. (n.d.). The publishing cycle [digital image]. https://library.uwinnipeg.ca/scholarly-communication/index.html

Usability.gov. (2021). User experience basics. *What & Why of Usability.* https://www.usability.gov/what-and-why/user-experience.html

Vaught, M. D., Jordan, D. C., & Bastian, H. (2016). *A cross-sectional study of commenters and commenting in PubMed, 2014 to 2016: Who's Who in PubMed Commons.* 9th International Congress on Peer Review and Scientific Publication. https://peerreviewcongress.org/abstract/a-cross-sectional-study-of-commenters-and-commenting-in-pubmed-2014-to-2016-whos-who-in-pubmed-commons/

Vaught, M., Jordan, D. C., & Bastian, H. (2017). Concern noted: A descriptive study of editorial expressions of concern in PubMed and PubMed Central. *Research Integrity and Peer Review, 2*(1), 1-11.

Vaught, M. (2017, July 12). PubMed Commons for journal clubs and authors [video]. *NCBI Minute webinar.* https://www.youtube.com/watch?v=dpM1S4XLgak&t=44s

Vaught, M., Jordan, D. C., & Bastian, H. (2019). PubMed Commons post-publication projects. *OSF Home.* https://osf.io/8xbqy/

Wakefield, A. J., … Walker-Smith, J. A. (1998). Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *The Lancet, 351,* 9103, 637-641.

Wakefield, A. J., Harvey, P., & Linnell, J. (2004). MMR—responding to retraction. *The Lancet, 363*(9417), 1327-1328.

Wang, Z., Sadovnick, A. D., Travoulsee, A. L., Rose, J. P., Bernales, C. Q., Encarnacion, M., …. Vilarino-Güell, C. (2016). Editorial note to: Nueclear receptor NR1H3 in familial multiple sclerosis. *Neuron, 92*(2), 331-332. doi: 10.1016/j.neuron.2016.10.008.

Wasko, M, & Faraj, S. (2005). Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. *MIS Quarterly, 29*(1), 35-57.

Wassermann, S., & Faust, K. (1994). *Social network analysis: Method and applications.* Cambridge University Press.

Watts, D. J., & and Stogatz, S. H. (1998). Collective dynamics of ‘small world’ networks. *Nature, 393*(6684). 409-410.

Web of Science. (n.d.). *Publons.* https://publons.com/about/mission

Weinberger, D. (2007). The real difference between the two 2.0s. *KM World, 16*(2), 18,20.

Wellman, B. (1982). Studying personal communities. In P. Marsden & N. Lin (Eds.), *Social Structure and Network Analysis* (pp. 61–80). Sage Publishing, Inc.

Wellman, B., & Gulia, M. (1999). Virtual communities as communities: Net surfers don’t ride alone. In M. A. Smith, & P. Kollock (Eds.), *Communities in cyberspace* (pp. 167-194). Routledge.

Wenger, E. (1998). *Communities of Practice.* Oxford University Press.

Wenger, E. (2000). Communities of Practice and social learning theory. *Organization, 7*(2), 225-246.

Wenger, E., Trayner, B., & De Laat, M. (2011). Promoting and assessing value creation in communities and networks: A conceptual framework. *Ruud de Moor Centrum.* http://www.knowledge-architecture.com/downloads/Wenger_Trayner_DeLaat_Value_creation.pdf

Weyhrich, S. (2013). *Sophistication & simplicity: The life and times of the Apple II computer.* Variant Press.

Williamson, P. O., & Minter, C. I. J. (2019). Exploring PubMed as a reliable resource for scholarly communications services. *Journal of the Medical Library Association, 107*(1), 16-29.

Wilson, T. D. (2000). Human Information Behavior. *Information Science Research, 3*(2), 49-55.

Xun, J., & Reynolds, J. (2010). Applying netnography to market research: The case of the online forum. *Journal of Targeting, Measurement and Analysis for Marketing, 18*(1), 17-31.

Yan, G., Tsekenis, G., Barzel, B., Slotine, J-J., Liu, Y-Y, Barabasi, A-L. (2015). Spectrum of controlling and observing complex networks. *Nature Physics, 11*, 779-786.

Yin, R. K. (2009). *Case study research: Design and methods (5th ed.).* Sage Publishing, Inc.

Yoon, S., Shin, J., Kim, S., Park, S., & Lee, J. B. (2012). Subject-based extraction of a latent blog community. *Information Sciences, 184*, 215-229.

Zhang, W., & Watts, S. (2004). Knowledge adoption in online communities of practice. *Systemes d'Information et Management, 9*(1), 81-102.

Zhang, W. (2007). Trust in electronic networks of practice: An integrative model. In Steinfield, C., Pentland, B. T., Ackerman, M., Contractor, N. (Eds.) *Communities and technologies 2007*. (pp. 351-370). https://10.1007/978-1-84628-905-7_18