

The impact of institutional repositories: a systematic review

Michelle R. Demetres; Diana Delgado, AHIP; Drew N. Wright

See end of article for authors' affiliations.

Objective: Institutional repositories are platforms for presenting and publicizing scholarly output that might not be suitable to publish in a peer-reviewed journal or that must meet open access requirements. However, there are many challenges associated with their launch and up-keep. The objective of this systematic review was to define the impacts of institutional repositories (IRs) on an academic institution, thus justifying their implementation and/or maintenance.

Methods: A comprehensive literature search was performed in the following databases: Ovid MEDLINE, Ovid EMBASE, the Cochrane Library (Wiley), ERIC (ProQuest), Web of Science (Core Collection), Scopus (Elsevier), and Library, Information Science & Technology Abstracts (EBSCO). A total of 6,593 citations were screened against predefined inclusion and exclusion criteria.

Results: Thirteen included studies were divided into 3 areas of impact: citation count, exposure or presence, and administrative impact. Those focusing on citation count (n=5) and exposure or presence (n=7) demonstrated positive impacts of IRs on institutions and researchers. One study focusing on administrative benefit demonstrated the utility of IRs in automated population of ORCID profiles.

Conclusion: Based on the available literature, IRs appear to have a positive impact on citation count, exposure or presence, and administrative burden. To draw stronger conclusions, more and higher-quality studies are needed.



See end of article for supplemental content.

INTRODUCTION

Gibbons defines an institutional repository (IR) as having the following core features: digital, community-driven and focused, institutionally supported, durable and permanent, and accessible [1]. These qualities make an IR an ideal platform for presenting and publicizing scholarly output that might not be suitable for publication in a peer-reviewed journal or that must meet open access (OA) requirements. This can include, but is not limited to, student work, presentations, working papers, conference papers, newsletters, electronic theses and dissertations (ETDs), journals with limited distribution, or electronic archival materials.

In creating an OA platform to showcase an institution's scholarly products, the benefits would seem to be self-evident, as the "OA advantage" in the traditional publishing environment has been widely discussed [2-4].

However, the challenges to developing an IR are varied and well documented [5-7]. Storage and staffing costs, low usage, faculty reticence to deposit in IRs, and time all align as reasons against the implementation and continued development of an IR at an academic institution. This systematic review aimed to define the various impacts that an IR can provide for an academic institution, thus justifying its implementation or maintenance.

METHODS

This study was performed following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [8]. In adherence to these guidelines, a protocol was registered in PROSPERO, an international prospective register of systematic reviews (registration # CRD42018091449).

Search strategy

Medical librarians performed comprehensive literature searches to identify studies that evaluated the impact of IRs on academic institutions. Initial searches were run on April 27, 2018, with an updated search on March 4, 2019. The following databases were searched: Ovid MEDLINE (ALL: 1946 to present), Ovid EMBASE (1974 to present), the Cochrane Library (Wiley), ERIC (ProQuest), Web of Science (Core Collection), Scopus (Elsevier), and Library, Information Science & Technology Abstracts (EBSCO). Google Scholar was not searched because of its inherent lack of reproducibility and unclear indexing practices. Search terms included all subject headings and associated keywords for “institutional repository,” “open access publishing,” “pre-print repository,” or “academic repository.” Specific IR names, derived from OpenDOAR, were also used as search terms [9]. The full search strategy for Ovid MEDLINE is available in the supplemental appendix. There were no language, publication date, or article type restrictions on the search strategy.

Study selection

After results were de-duplicated, 2 scholarly communications librarians independently screened a total of 6,593 citations using Covidence, a systematic review tool [10]. Discrepancies were resolved by consensus. Titles and abstracts were reviewed against predefined inclusion and exclusion criteria. Articles considered for inclusion were those that discussed demonstrated, measurable, or quantitative impacts of IRs. In light of the vast differences among repository types and content, the authors chose to focus only on IRs that were affiliated with academic institutions. For the purposes of this study, an academic institution was defined as an institution dedicated to education and research that grants academic degrees. An IR from an academic institution was defined as a web-based repository that exclusively supported content

created by members of said academic institution. Excluded studies were those that focused on non-academic institutions or OA platforms that were not institutionally affiliated. Excluded studies also included those that only discussed assumed or eventual impacts or benefits.

Full text was then pulled for selected studies for a second round of eligibility screening. Reference lists and citing articles for the studies selected for inclusion were also pulled and searched. A total of thirteen studies were selected for inclusion in this review. Figure 1 shows the full PRISMA flow diagram outlining the study selection process.

Data collection

Data from all included studies were pulled by two independent reviewers, assessing study design and outcomes, with results confirmed by consensus. After all included studies were reviewed, three areas of impact emerged: citation impact, exposure or presence, and administrative impact. Risk of bias was assessed at the individual study level according to standards set by the Cochrane Collaboration Qualitative Methods Group [11]. Per the standards, assessment of study quality included (i) adequacy of reporting detail, (ii) technical rigor and methodological soundness, and (iii) paradigmatic sufficiency. Due to the heterogeneity of the studies, no synthesis of results was performed.

RESULTS

The thirteen studies were divided into three areas of impact: citation count, exposure or presence, and administrative impact. Table 1 describes all included studies.

Citation count

Five of the thirteen included studies described the positive impact of IRs on citation count [12–16]. The most compelling of these, Gargouri 2010, demonstrated that the OA status of a paper resulting from its deposition in an IR was a statistically significant, independent, positive predictor of citation count, even when controlling for many other salient variables such as article age, journal impact factor, number of authors, number of pages, number of references cited, type of article, classification as a scientific article, and whether the first author was from the United States [14].

Figure 1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram

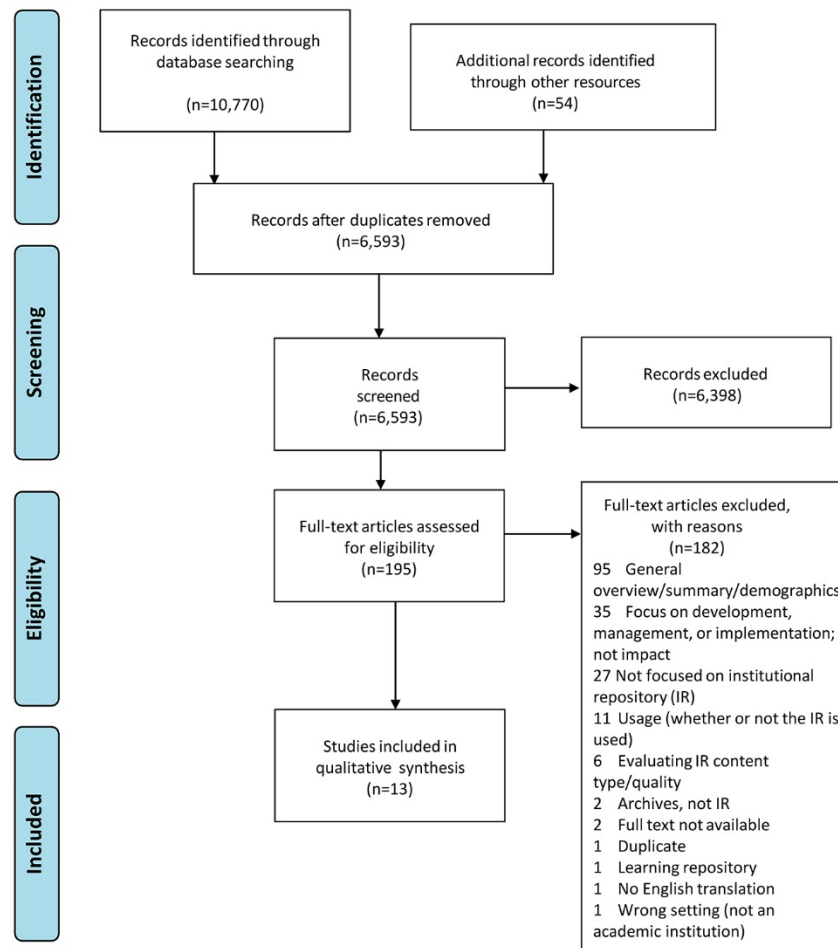


Table 1 Included studies

Study	Impact	Objective	Outcome	Quality
Atchison 2015 [13]	Citation count	Google Scholar was used to track citations and availability of self-archived papers.	Self-archived papers had more citations.	Low
Baessa 2015 [24]	Administrative	Institutionally affiliated authors were allowed to push publication information from the institutional repository (IR) into their ORCID profiles.	The IR populated and maintained up-to-date ORCID author profiles.	Moderate
Bangani 2018 [12]	Citation count	Citation counts were tracked for IR electronic theses and dissertations (ETDs). Citation analysis was then done on any journal article that was identified as resulting from these ETDs. Altmetrics (portable document format [PDF] views) were also documented.	Theses citations increased with the digitization of ETDs.	Moderate

Table 1 Included studies (continued)

Study	Impact	Objective	Outcome	Quality
Bruns 2014 [17]	Exposure or presence	Download statistics for master's theses were examined.	Thesis downloads from IR outpaced downloads from WorldCat.	Very low
Fan 2015 [18]	Exposure or presence	The contribution of IRs to their home institutions was calculated in terms of 4 webometric indicators: page counts, PDF counts, uniform resource locator (URL) mention counts, and link counts.	IRs improved webometric indicators of home institutions.	Low
Gargouri 2010 [14]	Citation count	Citation counts were compared between IR-deposited open access (OA) and non-OA articles published in the same (non-OA) journals.	OA due to deposition in IRs results in more citations.	Moderate
Linde 2012 [19]	Exposure or presence	The availability of conference proceedings stored in 6 IRs was examined.	25% of conference proceedings examined were only found in an IR.	Low
Organ 2006 [20]	Exposure or presence	Download statistics, page views, and cover views were tracked for an IR.	Materials in the IR were discoverable via Google more quickly than traditional publishing; downloads primarily came from Google.	Low
Pitol 2014 [15]	Citation count	Citation counts were collected via Google Scholar from an ~1,000-paper sample from 3 institutions.	Depositing in an IR, in combination with a listing in PubMed, resulted in more citations.	Low
Smith 2011 [21]	Exposure or presence	Internal links generated via Yahoo to IRs were traced back to Wikipedia.	Theses in IRs were used as evidence for Wikipedia articles.	Low
Smith 2013 [16]	Citation count	Deposit ratios of IRs with URL citation internal link counts were compared.	IRs with higher deposition rates were associated with more citations of their content.	Low
Stone 2014 [22]	Exposure or presence	Citations for ETDs in 49 IRs were tracked via Google Scholar.	ETDs were cited in peer-reviewed journals.	Low
Van Wyk 2014 [23]	Exposure or presence	Usage statistics of materials in IRs based on geographical location were evaluated.	IRs enhanced access to the global research community.	Low

ETDs in IRs also appeared to benefit from an OA advantage. Bangani 2018 tracked citation counts obtained through Google Scholar for ETDs published from 1989 to 2014 from North-West University in South Africa and found a marked increase in citations after ETDs began to be digitized. A total of 612 ETDs had 931 citations, translating to 1.52 citations per ETD on average. Prior to digitization, however, a total of 81 theses

and dissertations had only 10, translating to only 0.12 citations per document on average [12]. This positive citation impact pertained not only to the ETDs themselves, but also any published papers resulting from them. Pitol 2014 consistently found that IR deposition was the most impactful author-controlled way to make papers freely available in Google Scholar searches [15]. In combination with a listing in PubMed, IR deposition of papers resulted

in significantly more citations than papers that did not have freely available full text that could be found via Google Scholar.

Exposure or presence

Seven of the thirteen included studies focused on the increased exposure or discoverability that IRs provided for their content and their institutions [17–23]. This was measured in a variety of ways, including download statistics, webometric indicators (e.g., page views and internal links), and availability.

Linde 2012 saw increased exposure through the uniqueness of their collection: of the papers examined, 25% were not found digitally anywhere but the IR [19]. Van Wyk 2014 also spoke to the uniqueness of their collection at the University of Zululand, where exposure for their institutional output increased through the IR's overseas usage, which accounted for 5% of its overall usage [23]. Organ 2006 discussed the speed with which the IR made discoverability possible. Of the 80.9% of downloads coming from Google, most materials were discoverable by Google within 24–48 hours [20]. Bruns 2014 saw that a thesis, after being made available in the IR for 1 year, was downloaded 729 times. Approximately a year after that, the thesis was downloaded well over 3,500 times. Prior to IR deposition, this same thesis had only been downloaded 35 times during a 1-year period [17].

Administrative

Baessa 2015 discussed the unique administrative benefit provided by their IR [24]. The King Abdullah University of Science and Technology (KAUST) leveraged the benefit of the IR by allowing KAUST-affiliated authors to push publication information from the IR into their ORCID profiles. This helped authors maintain a current public profile without having to manually update their profiles themselves.

DISCUSSION

One limitation of this review was that, for the most part, the quality of the articles discussing the impact of IRs was poor. Few discussed measurable or quantitative benefits, whereas many discussed the overall development of their IRs and their anticipated impact. Methodology was generally poor as well, with small sample sizes and questionable rationales. Therefore, more

quantitative, methodologically rigorous studies are needed in this area.

The goal of making underrepresented work, such as ETDs, discoverable and citable can be achieved through making this work available in an IR and discoverable by Google and Google Scholar, which is also often linked to an improved citation rate. As Stone 2014 points out, “many senior theses are about regional and local issues or cutting-edge topics, both of which may have a dearth of publications in the mainstream scholarly literature due to interest and/or the scholarly publishing cycle” [22]. This makes an IR an important source for studies on new or niche topics and an accessible avenue for scholars.

Indexing and discoverability are the main reasons for the difference in the “OA advantage” (e.g., more citations) between IRs and OA journals. OA journals are often widely indexed in databases like MEDLINE, EMBASE, and Scopus, which makes discoverability for OA journal articles, and therefore citations, much more likely. IRs are typically only discoverable through Google and Google Scholar or OA content aggregators such as CORE [25]. The Institutional Repository LinkOut feature in PubMed is still not widely adopted, with only 36 IRs currently participating [26]. This is likely due to the difficult requirements for IR application and acceptance (e.g., minimum of 1,000 articles that are not already deposited in PubMed Central) [27].

However, the literature suggests that IRs can still be an important outlet for exposure, particularly when used for preprints and data. Conroy found “journal articles that were uploaded as preprints before being published gather more citations in the long run than papers without a preprint version” [28]. The effect is similar with regard to the availability of data. One study found that articles with data availability statements have up to 25% higher citation impact on average [29]. Another found that publicly available data are associated with a “69% increase in citations, independently of journal impact factor, date of publication, and author country of origin” [30]. Most articles included in this review focused on ETDs or traditional journal articles. However, IRs are also important potential platforms for the exposure and citation of preprints and data.

One justification for IRs comes from the Linde 2012 study, which reports that 25% of conference

proceedings examined were only found in the IR [19]. The inaccessibility of conference or meeting abstracts is a problem frequently confronting librarians and researchers, especially those who seek to include grey literature in systematic reviews. Not all conferences publish proceedings, effectively making this work inaccessible. For those that do, the representation of the work is often incomplete or impermanent. For example, poster presentations are frequently published as written abstracts only, omitting the context and graphics that would be included in a physical poster. This is not the case with IRs, in which complete posters can be made available.

Moreover, conference and meeting abstracts are usually published as supplements, for which online access is not always guaranteed in perpetuity. Finding full text for older abstracts is often impossible. If published on association websites, the abstracts may be made available only to attendees of the specific meeting or members of the organization [31]. Furthermore, conference and meeting abstracts often do not result in full-length publications [32–34], with one study showing that of a total of 29,729 abstracts presented at scientific meetings, the rate of weighted mean abstracts that were fully published was 44.5% [35]. This suggests that the conference and meeting abstracts are the only source of a large portion of scientific information. Together, these reasons make a compelling argument for using IRs as outlets for conference and meeting abstracts and posters.

The studies included in this review also provide evidence of the global impact of IRs. Six of the thirteen included studies came from institutions outside the United States [12, 18–20, 23, 24]. This speaks to the interconnectivity that IRs can facilitate, making scholarship available internationally. Much discussion in scholarly communication circles focuses on the disparity in access to and production of research between the Global North and Global South [36–38]. As Vattikoti points out, “Most of the countries in Global south are not in a position to afford such huge fees charged by pay-access publishers because of insufficient funds or prioritization of the limited research funds for carrying out research activities” [39]. By contributing to a free-to-deposit OA platform like an IR, researchers help bridge this scholarship gap and position IRs as an important resource in equitizing academic scholarship.

Based on the available literature, the authors found that IRs appear to have a positive impact on citation count, exposure or presence, and administrative burden. To make stronger conclusions, more and higher-quality studies are needed.

REFERENCES

- Gibbons S. Chapter 2: defining an institutional repository. *Libr Technol Rep*. 2004 Jun 2;4:6–10.
- Eysenbach G. Citation advantage of open access articles. *PLoS Biol*. 2006 May 16;4(5):e157.
- Breugelmans JG, Roberge G, Tippet C, Durning M, Struck DB, Makanga MM. Scientific impact increases when researchers publish in open access and international collaboration: a bibliometric analysis on poverty-related disease papers. *PLoS One*. 2018 Sep 19;13(9):e0203156.
- Piwowar H, Priem J, Larivière V, Alperin JP, Matthias L, Norlander B, Farley A, West J, Haustein S. The state of OA: a large-scale analysis of the prevalence and impact of open access articles. *Peer J*. 2018 Feb 13;6:e4375.
- Allard S, Mack TR, Feltner-Reichert M. The librarian’ role in institutional repositories. *Ref Serv Rev*. 2005 Sep;33(3):325–6.
- Buehler MA, Boateng A. The evolving impact of institutional repositories on reference librarians. *Ref Serv Rev*. 2005 Sep;33(3):291–300.
- Foster NF, Gibbons S. Understanding faculty to improve content recruitment for institutional repositories. *D-Lib Mag*. 2005 Jan;11(01).
- Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, Clarke M, Devereaux PJ, Kleijnen J, Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med*. 2009 Jul 21;6(7):e1000100.
- University of Nottingham. OpenDOAR: directory of open access repositories [Internet]. The University [cited 14 Mar 2018]. <<http://www.openoar.org/>>.
- Covidence. Covidence systematic review software [Internet]. Melbourne, Australia: Veritas Health Innovation; 2019 [cited 5 Sep 2019]. <<http://www.covidence.org/>>.
- Hannes K. Chapter 4: Critical appraisal of qualitative research. In: Noyes J, Booth A, Hannes K, Harden A, Harris J, Lewin S, Lockwood C, eds. *Supplementary guidance for inclusion of qualitative research in Cochrane systematic reviews of interventions*. Cochrane Collaboration Qualitative Methods Group; 2011.
- Bangani S. The impact of electronic theses and dissertations: a study of the institutional repository of a university in South Africa. *Scientometrics*. 2018 Apr;115(1):131–51.
- Atchison A, Bull J. Will open access get me cited? an analysis of the efficacy of open access publishing in political science. *Polit Sci Polit*. 2015 Jan;48(01):129–37.

14. Gargouri Y, Hajjem C, Larivière V, Gingras Y, Carr L, Brody T, Harnad S. Self-selected or mandated, open access increases citation impact for higher quality research. *PLoS ONE*. 2010 Oct 18;5(10):e13636.
15. Pitol SP, De Groot SL. Google Scholar versions: do more versions of an article mean greater impact? *Libr Hi Tech*. 2014 Nov 11;32(4):594–611.
16. Smith AG. Web based impact measures for institutional repositories [Internet]. 2013;1806–16 [cited 5 Sep 2019]. <<https://core.ac.uk/download/pdf/41338171.pdf>>.
17. Bruns TA, Knight-Davis S, Corrigan EK, Brantley S. It takes a library: growing a robust institutional repository in two years. *Coll Undergrad Libr*. 2014 Jul 3;21(3–4):244–62.
18. Fan W. Contribution of the institutional repositories of the Chinese Academy of Sciences to the webometric indicators of their home institutions. *Scientometrics*. 2015 Dec;105(3):1889–909.
19. Linde P, Eriksson J, Kullman L, Fathli M, Karlsson K, Sikström M, Sköld Y, Tång I. Accessibility and self-archiving of conference articles: a study on a selection of Swedish institutional repositories. *Inf Serv Use*. 2012 Oct 16;31(3–4):259–69.
20. Organ M. Download statistics: what do they tell us? *D-Lib Mag*. 2006 Nov;12(11).
21. Smith AG. Wikipedia and institutional repositories: an academic symbiosis? *Victoria University of Wellington*; 2011. p. 794–800.
22. Stone SM, Lowe MS. Who is citing undergraduate theses in institutional digital repositories? implications for scholarship and information literacy. *Coll Undergrad Libr*. 2014 Jul 3;21(3–4):345–59.
23. Van Wyk B, Mostert J. African institutional repositories as contributors to global information: a South African case study. *Mousaion*. 2014;32(1):98–114.
24. Baessa M, Lery T, Grenz D, Vijayakumar JK. Connecting the pieces: using ORCIDs to improve research impact and repositories. [version 1; peer review: 2 approved]. *F1000 Res*. 2015 Jul 7;4:195.
25. The Open University, Jisc. CORE: the world's largest collection of open access research papers [Internet]. The University [cited 22 Oct 2019]. <<https://core.ac.uk/>>.
26. National Center for Biotechnology Information, National Library of Medicine. Other LinkOut resources by SubjectType: institutional repository [Internet]. The Library [cited 22 Oct 2019]. <https://www.ncbi.nlm.nih.gov/projects/linkout/journals/htmllists.cgi?type_id=16#institutional%20repository>.
27. National Center for Biotechnology Information, National Library of Medicine. How do institutional repositories (IR) join LinkOut to display links to the IR in PubMed? [Internet]. The Library [cited 22 Oct 2019]. <<https://www.ncbi.nlm.nih.gov/projects/linkout/doc/IR-application.shtml>>.
28. Conroy G. Preprints boost article citations and mentions. *Nature Index*. 2019 Jul 9.
29. Colavizza G, Hrynaszkiewicz I, Staden I, Whitaker K, McGillivray B. The citation advantage of linking publications to research data. *arXiv*. 2019;1907.02565v1.
30. Piwowar HA, Day RS, Fridsma DB. Sharing detailed research data is associated with increased citation rate. *PLoS ONE*. 2007 Mar 21;2(3):e308.
31. Tulane University Libraries. Conference proceedings and meeting abstracts in the health science literature: a guide: how to locate conference proceedings [Internet]. The Libraries; 2019 [cited 12 Sep 2019]. <<https://libguides.tulane.edu/proceedings>>.
32. Klassen TP, Wiebe N, Russell K, Stevens K, Hartling L, Craig WR, Moher D. Abstracts of randomized controlled trials presented at the Society for Pediatric Research meeting: an example of publication bias. *Arch Pediatr Adolesc Med*. 2002 May;156(5):474–9.
33. Shelmerdine SC, Lynch JO, Langan D, Arthurs OJ. Presentation to publication: proportion of abstracts published for ESPR, SPR and IPR. *Pediatr Radiol*. 2016 Sep;46(10):1371–7.
34. Bhandari M, Devereaux PJ, Guyatt GH, Cook DJ, Swiontkowski MF, Sprague S, Schemitsch EH. An observational study of orthopaedic abstracts and subsequent full-text publications. *J Bone Joint Surg Am*. 2002 Apr;84(4):615–21.
35. Scherer RW, Langenberg P, von Elm E. Full publication of results initially presented in abstracts. *Cochrane Database Syst Rev*. 2007 Apr 18;(2):MR000005.
36. Cash-Gibson L, Rojas-Gualdrón DF, Pericàs JM, Benach J. Inequalities in global health inequalities research: a 50-year bibliometric analysis (1966–2015). *PLoS ONE*. 2018 Jan 31;13(1):e0191901.
37. Raju R, Pietersen J. Library as publisher: from an African lens. *J Electron Publ*. 2017 Aug 1;20(2).
38. Open Library of Humanities. Open insights: an interview with Leslie Chan [Internet]. The Library; 2018 [cited 12 Sep 2019]. <<https://www.openlibhums.org/news/314/>>.
39. Vattikoti K. Evolution of open access publishing towards a more equitable solution. *J Legal Ethic Regul Issues*. 2019;22(3).



SUPPLEMENTAL FILE

- **Appendix:** [The full search strategy for Ovid MEDLINE](#)

AUTHORS' AFFILIATIONS

Michelle R. Demetres, mrd2006@med.cornell.edu,
<http://orcid.org/0000-0002-4997-7707>, Scholarly Communications
Librarian, Samuel J. Wood Library, Weill Cornell Medicine, New York,
NY

Diana Delgado, AHIP, did2005@med.cornell.edu,
<https://orcid.org/0000-0002-6290-3497>, Associate Director,
Information, Education and Clinical Services, Samuel J. Wood Library
& C.V. Starr Biomedical Information Center, Weill Cornell Medicine,
New York, NY

Drew N. Wright, drw2004@med.cornell.edu, <https://orcid.org/0000-0002-1776-5427>, Research Librarian, Weill Cornell Medicine, New York, NY

Received September 2019; accepted December 2019



Articles in this journal are licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).



This journal is published by the [University Library System](#) of the [University of Pittsburgh](#) as part of its [D-Scribe Digital Publishing Program](#) and is cosponsored by the [University of Pittsburgh Press](#).

ISSN 1558-9439 (Online)