Research Article

Burden and Characteristics of Unsolicited Emails from Medical/Scientific Journals, Conferences, and Webinars to Faculty and Trainees at an Academic Pathology Department

Matthew D. Krasowski[1], Janna C. Lawrence[2], Angela S. Briggs[1], Bradley A. Ford[1]

[1]Department of Pathology, University of Iowa Hospitals and Clinics, Iowa City, IA, USA, [2]Hardin Library for The Health Sciences, University of Iowa, Iowa City, IA, USA

Received: 04 February 2019 Accepted: 07 March 2019 Published: 06 May 2019

Abstract

Background: Professionals and trainees in the medical and scientific fields may receive high e-mail volumes for conferences and journals. In this report, we analyze the amount and characteristics of unsolicited e-mails for journals, conferences, and webinars received by faculty and trainees in a pathology department at an academic medical center. **Methods:** With informed consent, we analyzed 7 consecutive days of e-mails from faculty and trainees who voluntarily participated in the study and saved unsolicited e-mails from their institutional e-mail address (including junk e-mail folder) for medical/scientific journals, conferences, and webinars. All e-mails were examined for characteristics such as reply receipts, domain name, and spam likelihood. Journal e-mails were specifically analyzed for claims in the message body (for example, peer review, indexing in databases/resources, rapid publication) and actual inclusion in recognized journal databases/resources. **Results:** A total of 17 faculty (4 assistant, 4 associate, and 9 full professors) and 9 trainees (5 medical students, 2 pathology residents, and 2 pathology fellows) completed the study. A total of 755 e-mails met study criteria (417 e-mails from 328 unique journals, 244 for conferences, and 94 for webinars). Overall, 44.4% of e-mails were flagged as potential spam by the institutional default settings, and 13.8% requested reply receipts. The highest burden of e-mails in 7 days was by associate and full professors (maximum 158 or approximately 8200 per year), although some trainees and assistant professors had over 30 e-mails in 7 days (approximately 1560 per year). Common characteristics of journal e-mails were mention of "peer review" in the message body and low rates of inclusion in recognized journal databases/resources, with 76.4% not found in any of 9 journal databases/resources. The location for conferences in e-mails included 31 different countries, with the most common being the United States (33.2%), Italy (9.8%), China (4.9%), United Kingdom (4.9%), and Canada (4.5%). **Conclusions:** The present study in an academic pathology department shows a high burden of unsolicited e-mails for medical/scientific journals, conferences, and webinars, especially to associate and full professors. We also demonstrate that some pathology trainees and junior faculty are receiving an estimated 1500 unsolicited e-mails per year.

Keywords: Electronic mail, open access publishing, peer review, publishing, scholarly communication, scientific conferences, webinars

Introduction

Peer-reviewed journal publications and conference presentations are critical activities for those in academic and commercial research, impacting admission to schools, consideration for awards, and promotion. The last two decades have seen an explosion of medical and scientific journals, many operating solely online. Some of these journals have been criticized for their questionable quality (e.g., absent or minimal peer review), unclear editorial oversight, and mass e-mail marketing tactics.[1-9] One study showed acceptance of an intentionally constructed "spoof" article containing bogus data by a high percentage of open access journals.[10,11] The term "predatory journal" has been used to denote journals of

Access this article online

Quick Response Code:

Website: www.jpathinformatics.org

DOI: 10.4103/jpi.jpi_12_19

Address for correspondence: Dr. Matthew D. Krasowski, Department of Pathology, University of Iowa Hospitals and Clinics, 200 Hawkins Drive, C-671 GH, Iowa City, IA 52242, USA. E-mail: mkrasows@healthcare.uiowa.edu

How to cite this article: Krasowski MD, Lawrence JC, Briggs AS, Ford BA. Burden and characteristics of unsolicited emails from medical/scientific journals, conferences, and webinars to faculty and trainees at an academic pathology department. J Pathol Inform 2019;10:16.

Available FREE in open access from: http://www.jpathinformatics.org/text.asp?2019/10/1/16/257736

poor quality that often use aggressive marketing, although a standard definition for this term has not yet been agreed upon.

Analogous to journals, concern has also arisen regarding conferences or meetings that are of potentially low quality and scientific value (sometimes termed "fraudulent" or "predatory" conferences).[3,12-14] There has been less analysis of these conferences compared to journals, but common themes include organization by companies or individuals as opposed to established scientific communities or organizations, high fees for presenters, and lack of transparency regarding the details of the meeting. Such conferences may be in a variety of countries and solicit researchers for a variety of conference activities. In addition to conferences and meetings, researchers may also receive solicitations for webinars that charge for access and have unclear quality.

Those involved in medical and scientific professions often receive unsolicited e-mails related to journals, conferences, and webinars. The e-mails may include invitations to submit articles or join editorial boards for online publications. For conferences, the e-mails may be invitations to present data, serve on speaker panels, or to organize conference sessions. One early career researcher reported over 500 e-mails from journals and conferences in the year following his first publication as corresponding author.[13] The majority of invitations were for journals and conferences outside his field of interest.

The present study analyzed volume and characteristics of unsolicited e-mails related to medical/scientific journals, conferences, and webinars sent to faculty and trainees at an academic pathology department over a 7-day period. Variables analyzed for all e-mails included whether the message was recognized as potential spam by the institutional e-mail settings, request of read receipts, and domain name. E-mails for journals were analyzed for claims by the sender such as peer review, inclusion in databases/resources, and rapid peer review/publication. The actual inclusion of journals in recognized databases/resources was ascertained. The data across academic rank of the participants were also compared.

Methods

Setting

The study setting was a pathology department in an academic medical center. Trainees of the department include externs (medical students taking a year to work in pathology in the middle of medical school), pathology residents, and pathology fellows. Faculty are either on clinical or tenure tracks. The study had institutional review board approval and required informed consent from all participants. For purposes of analysis, pathology residents and fellows were combined. Clinical and tenure-track faculty at the same rank (i.e., assistant, associate, or full professor) were combined.

Study design

The study design was as follows. All faculty and trainees in the department were sent an e-mail describing the premise of the study and with a link to a short online questionnaire which provided more detailed information. The e-mail emphasized that participation was voluntary. For those who were eligible and interested in participating further, directions were provided on how to save 7 consecutive days of unsolicited e-mails from journals (including invitations for editorships), scientific/medical conferences, and webinars that had been sent to their institutional e-mail address, both to the main inbox and spam (junk mail) folders. The institution uses Microsoft Outlook as the primary e-mail platform. Study participants saved unsolicited e-mails in Outlook folders and then exported these as Outlook ".pst" files to a network folder created for the study. The survey included a question on whether participants actively manage their spam filters (i.e., alter them from the institutional default settings by measures such as blocking specific sender domains) to limit unsolicited e-mails.

Analysis of e-mails

Using the analysis described below [computer code provided in Supplement File 1], the following information/data was extracted from the e-mails: subject, from name, from address, reply to address, received date/time, message body, flag indicating whether reply receipt had been requested, domain name, and Microsoft Exchange Spam Confidence Level (SCL).[15] Processing of Outlook ".pst" files was scripted at the command line in Linux (Ubuntu 18.04.1 LTS). The "readpst" function of the libpst utility (version 0.6.72; https://www.five-ten-sg.com/libpst/) was first used to convert individual files to nonproprietary mbox format as follows:

readpst junkmail. pst > junkmail. mbox

mbox files were further parsed into comma-separated files that could be imported into Microsoft Excel for manual review. This was accomplished using modified open-source Python code [Supplement File 1A].

Data derived from ".pst" files using the code above are not human readable in all cases. For example, the spam score (SCL) used by Microsoft Outlook (embedded in the X-Forefront-Antispam-Report header extracted above) is complex but the numerical score can be extracted using regular expressions in Microsoft Excel. To enable this, Visual Basic code was added to a template spreadsheet into which individual comma-separated files were imported. Code to extract the SCL was called as follows (in this case, the X-Forefront-Antispam-Report header for a particular E-mail is in Excel cell D3, and an mbox file has been imported into a tab called "Data"):

=RegexExtract (Data!D3,"SCL[:](\d+)[;]")

The classification of mail as spam or not spam according to our local configuration was therefore readable as:

=RegexExtract (Data!D3,"SFV[:](\w+)[;]")

This code also facilitated comparison of the "from" and "reply to" addresses, as addresses could be extracted from between

"<>" angle brackets and compared with an excel formula as follows:

*=IF (RegexExtract (Data!B2,"[<](. *)[>]") =RegexExtract (Data!C2,"[<](. *)[>]"),1,0)*

The visual basic code is provided in Supplement File 1B.

Data analysis

These data were further analyzed in spreadsheets. Each e-mail was then manually reviewed to determine whether the e-mail fit the inclusion categories of the study and if so was then assigned to the categories of journals, conferences, and webinars. For e-mails involving journals, the following information was captured through a combination of spreadsheet text matches and manual review of the text of e-mail message: full journal name, mention of articles in the journal being peer-reviewed, mention of "rapid" in reference to peer review and/or speed of publication for the journal, mention of indexing or potential indexing in PubMed and/or MEDLINE (sometimes with indirect language such as "some journals have been indexed in PubMed or PubMed Central"), mention of being included in Index Copernicus, mention of having a numeric factor (such as impact or citation factor) other than the Index Copernicus Value in one or more journal databases/resources, and International Standard Serial Number (ISSN) number (to help verify journal name, if needed). The Index Copernicus Value was included in the variables as it was identified in a previous study as a metric commonly promoted by potentially predatory journals.[16] The journal name was verified by accessing the journal home page on the Internet.

E-mails related to conferences were manually reviewed to ascertain the country in which conference was being held, accessing the homepage for the conference on the Internet to verify details if not explicitly clear in the e-mail. E-mails related to webinars were reviewed to verify that the e-mail was for a webinar advertisement. For both conferences and webinars, manual review was also performed to make sure that the e-mails were advertisements and not a conference or webinar for which the e-mail recipient had registered for. We did not find any examples of this in the dataset (i.e., based on e-mail content, the e-mails for conferences and webinars selected by the participants appeared to be uniformly unsolicited).

A total of 823 e-mails from 26 participants (9 trainees, 17 faculty) were submitted by study participants. A total of 68 did not meet inclusion criteria, including 34 that were from newsletters and not journals. An additional 34 did not meet inclusion criteria for various other reasons such as being advertisements for medical or scientific products (e.g., laboratory reagents such as antibodies) or being related to a journal activity in which the e-mail recipient was actually participating in (e.g., an e-mail from journal editor/publisher related to actual involvement of the recipient such as submission or peer review of an article). After exclusions, this left 755 e-mails that form the basis for the analysis reported in this study.

Results

General characteristics of unsolicited e-mails

A total of 26 people completed the entire study including 5 medical students, 2 pathology residents, 2 pathology fellows, 4 assistant professors (3 clinical, 1 tenure track), 4 associate professors (all clinical track), and 9 full professors (5 clinical, 4 tenure track). Of the participants, 11 (42.3%) actively manage e-mail spam filter settings to block unsolicited e-mails. A total of 755 e-mails met the study criteria and included 417 e-mails from 328 unique journals, 244 e-mails for conferences, and 94 e-mails for webinars. For the journal-related e-mails, 94 (22.5%) were requests to be an editor or join editorial board; the remainder related to journal article submission. Figure 1 summarizes some general characteristics of the e-mails for journals, conferences, and webinars. Overall, 44.4% of e-mails were identified as potential spam using the default institutional settings diverting e-mails with SCL[15] of 5 or 6 to a junk mail folder [Figure 1a; note that we captured the SCL assigned to the e-mail message independent of any additional measures taken by users to limit or divert spam e-mail such as blocking specific domains]. Journal e-mails were identified as potential spam at 56.6% compared to 29.2% and 29.0%, respectively, for conference and webinar e-mails. Journals and conference e-mails requested reply receipts at 16.3% and 14.8%, respectively; in contrast, none of the 94 webinar e-mails requested reply receipts [Figure 1b]. Journals, conferences, and webinar e-mails had similar rates (31.2%–40.0%) of the "from address" matching "reply to address" [Figure 1c]. The most common top-level domain names for the e-mails were the generic domains com (58.6%), org (21.9%), net (7.7%), and info (5.5%) [Figure 1d]. The remainder was comprised of biz, edu, and country code top-level domains. Redirection of e-mails through intermediate hosts obscured the country and sender of origin in most cases; so, locations of conferences as well as journal names were derived through manual review.

Variation in e-mail burden by academic rank

We next analyzed how the burden of e-mail varied by academic rank. Five of the 7 medical students, residents, and fellows had <10 unsolicited e-mails in the entire week. The highest number of total e-mails were seen with associate professors (one faculty member with 68 e-mails) and full professors (3 participants with 67, 74, and 158 e-mails, respectively). Figure 2a aggregates the data for journal, conference, webinar, and total e-mails and averages by academic rank of the e-mail recipient. In this plot, the trend upward by rank is readily apparent. Webinar e-mails were uncommon outside of associate and full professors. Figure 2b shows the data subdivided by the number of journal articles published by the participant and included in PubMed (<10, 10–30, 31–100, or >100 articles), a surrogate for publishing productivity of the participant. As can be seen, the trends are very similar to the data analyzed by academic rank shown in Figure 2a. The estimated percent of total e-mails comprised by the unsolicited e-mails for each participant was averaged by

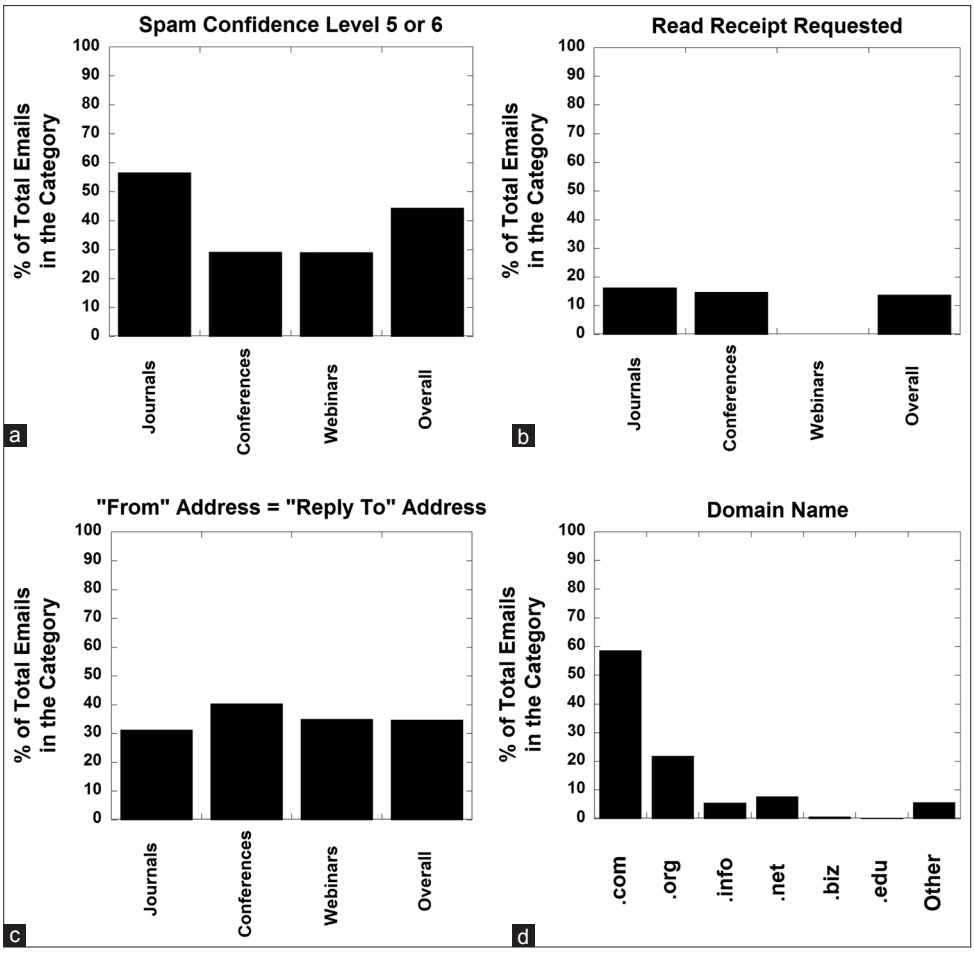

Figure 1: General characteristics of unsolicited e-mails for journals, conference, and webinars. The graphs detail percent of e-mails which met the following criteria: (a) Spam confidence level of 5 or 6, (b) request for read receipt, (c) "From address" matching "reply to" address, and (d) top-level domain names for e-mails

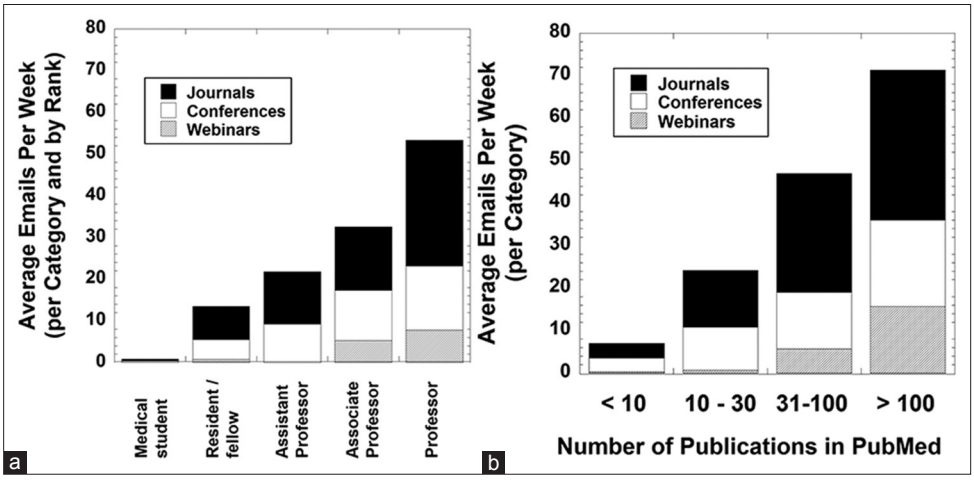

Figure 2: Volumes of e-mails by academic rank and publishing history. (a) E-mails aggregated by the academic rank of participant. (b) E-mails aggregated by categories for number of articles participant has been coauthor on that is included in PubMed

academic rank. The percent of total was similar for all academic rank categories (~7%–8%) except medical students (medical students, 1.2%; residents/fellows, 8.5%; assistant professors, 8.5%; associate professors, 7.3%; full professors, 8.1%).

Characteristics of journals in unsolicited e-mails

We reviewed all e-mails from journal publishers to identify characteristics of the e-mail message body. As summarized in Figure 3a, the most common claim in the e-mail message was

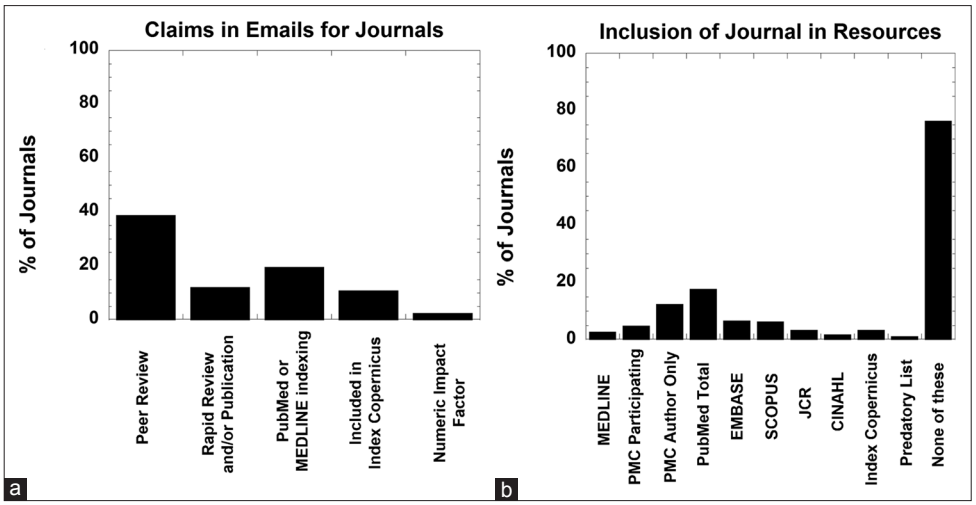

Figure 3: Claims and actual database indexing of journals in e-mails. (a) Percent of journal e-mails with message body claiming "peer review" of articles, rapid review and/or publication of submitted articles, indexing in PubMed and/or MEDLINE, inclusion in Index Copernicus, and a numeric impact factor other than the Index Copernicus IC Index. (b) Percent of unique journals from e-mails officially indexed or included in MEDLINE, PubMed, PubMed Central, EMBASE, Scopus, Journal Citation Reports, CINAHL, Index Copernicus, and Stop Predatory Journals. See Table 1 for details on these databases

of peer review of articles (38.8%), with other claims below 20%. Mention of a numeric factor (e.g., impact or citation factor) other than the Index Copernicus Value was found in only 2.6% of e-mails. Inclusion in Index Copernicus was mentioned in 11.0% of e-mails; some reference to the journal being indexed or included in PubMed, PubMed Central, and/or MEDLINE was in 19.7% of e-mails.

We next ascertained whether the journals were actually officially indexed or included in the following databases/resources: MEDLINE, PubMed, PubMed Central, EMBASE, Scopus, Journal Citation Reports (JCR), CINAHL, Directory of Open Access Journals (DOAJ), Index Copernicus, and "Stop Predatory Journals" website[17] [Table 1 has details on these sources]. For PubMed Central, we distinguished between those journals that routinely deposit articles into PubMed Central versus journals that appear in PubMed Central solely from author-initiated deposit of articles arising from National Institutes of Health (NIH) funding. The PubMed Central "Participating" journals include those that are listed on the PubMed Central database as full participation (journals that deposit complete contents in PubMed Central), NIH portfolio (journals that deposit all NIH-funded articles per NIH Public Access Policy and possibly additional non-NIH-funded articles), or selective deposit (journals that deposit a subset of articles and/or offer a hybrid open access model). The PubMed Central "Author Only" journals would otherwise not be included in PubMed Central if not for author-initiated deposits of NIH-funded research.[18] Inclusion of journals in databases/resources was checked at least 6 months after the e-mail receipt. This allowed for catching journals that were in the process of being added to databases/resources at the time of the e-mail.

Most journals (76.4%) were found in none of the databases/resources examined. A total of 17.7% of the journals were included in PubMed, with 12.5% as PubMed Central Author Only journals, 4.9% as PubMed Central Participating journals, and 2.7% in MEDLINE (note there were 6 journals that are both in MEDLINE and a PubMed Central Participating journal). The journals were found at a rate of 6.7% or less in all other databases/resources [Figure 3b]. The 41 journals in our dataset that were PubMed Central Author Only journals had a mean of only 4.2 articles (standard deviation, 7.5; median, 1; range 1–45) included in the entire PubMed database. These 41 journals had either a single article in PubMed or publications scattered across widely separated dates. In contrast, the 19 journals in our dataset that were indexed in MEDLINE and/or are a PubMed Central Participating journal have a mean of 1374 articles (standard deviation, 1692; median, 605; range 46–5,235) in the entire PubMed database. Interestingly, only 3.4% of the journals were actually found in the Index Copernicus World Journal Masters List and had an Index Copernicus Value despite 11.0% of journals claiming inclusion in Index Copernicus in the e-mail text.

Characteristics of unsolicited e-mails for webinars and conferences

A total of 94 e-mails were related to webinars, with 96.8% (91 of 94) received by associate and full professors, with weighting toward those with more publications of their own in PubMed [Figure 2a and b]. All e-mails in this category were solicitations to pay a fee to view a webinar; none were invitations for the e-mail recipient to be the presenter or organizer of the event. The remaining 3 were received by one pathology fellow; none were received by medical students, pathology residents, or assistant professors in the study. For three of the faculty, webinar e-mails accounted for 25.0% or more of their total unsolicated e-mails in the week. These three faculties each had 40 or more of their own publications in PubMed.

In contrast to webinars, unsolicited e-mails for conferences were more evenly spread across academic rank and recipient

Table 1: Journal databases/resources

Database	Approximate number of journals	Number of records	Maintained by	Comments	Link
CINAHL	5500	6,000,000	EBSCO	EBSCO publishes numerous resources; CINAHL focuses on nursing/allied health resources	https://health.ebsco.com/products/the-cinahl-database
EMBASE	8500	32,000,000	Elsevier	Elsevier is a publisher; EMBASE covers Medline plus~2000 other biomedical journals, mostly European. Also includes conference abstracts	https://www.embase.com/login
Index Copernicus	45,500 (6500 with Index Copernicus Value)	Not applicable	Index Copernicus International	Includes nonjournal data separately, with a focus on non-English-language journals and qualitatively defined numeric rankings	https://journals.indexcopernicus.com/
DOAJ	12,000	3,725,000	Infrastructure services for open access CIC	Directory of Open Access Journals is an independently curated not-for-profit membership-based database	https://doaj.org
Journal Citation Reports	11,500	2,200,000	Clarivate Analytics	Integrated with the subscription ISI Web of Science, source of proprietary JIF	https://clarivate.com/products/journal-citation-reports/
MEDLINE	5200	25,000,000	U.S. NLM	Primary component of PubMed, made available to commercial suppliers	https://www.nlm.nih.gov/bsd/medline.html
PubMed	30,000	29,000,000	U.S. NLM	Produced by the NLM and freely available. Includes MEDLINE and PubMed Central	https://www.ncbi.nlm.nih.gov/pubmed/
PubMed Central	7460	5,200,000	U.S. NLM	Subset of PubMed, number includes only full participation, NIH portfolio, and selective deposit journals; does not include journals with only author-deposited articles	https://www.ncbi.nlm.nih.gov/pmc/
Scopus	22,800	71,000,000	Elsevier	As with EMBASE, administered by Elsevier, but with an independent board governing content	https://www.elsevier.com/solutions/scopus
Stop Predatory Journals	1400	Not applicable	Anonymous	Lists potentially predatory journals; 10 criteria for defining "predatory" are published at the URL at right	https://predatoryjournals.com/about/

CIC: Community Interest Company, NLM: National Library of Medicine, JIF: Journal Impact Factor, NIH: National Institutes of Health, CINAHL: Cumulative Index to Nursing and Allied Health Literature

PubMed publications [Figure 2]. A total of 244 e-mails for conferences were received during the week, with 81 (33.2%) related to conferences held within the United States. The remaining 163 (66.8%) conferences were held in 30 additional countries, with the most common being Italy (9.8%), China (4.9%), United Kingdom (4.9%), Canada (4.5%), Spain (4.1%), Belgium (2.9%), Netherlands (2.9%), and Thailand (2.9%). All other countries were at 2.5% or below.

There were 3 primary purposes of the e-mails: invitation to attend the conference (34.8%), request for submission of abstracts or papers (27.5%), or solicitation to be speaker and/or organizer (37.7%). The last category was somewhat heterogeneous with some e-mails having combined requests (e.g., speaker or keynote speaker and platform organizer).

Discussion

The present study illustrates wide variability in the amount of unsolicited e-mails for journals, conferences, and webinars to the institutional e-mail addresses for faculty and trainees in an academic pathology department. At the low end, 5 of 7 trainees (medical students, residents, and fellows) received 10 or less e-mails in a week of analysis. At the other extreme are faculty who received up to 158 e-mails in 1 week or approximately 8200 per year. Six of the fellows, assistant professors, and associate professors had a similar number of total e-mails (~500) to those reported by an early career researcher who analyzed a year's worth of e-mails following his first publication as a corresponding author.[13] In all groups except medical students, these e-mails accounted for approximately 8% of total e-mails. The institutional spam filter default settings divert 44.4% of these e-mails to junk mail folders, assuming the user has not taken additional steps to block unsolicited e-mails. For those receiving the highest amounts of unsolicited e-mails, just deleting these messages consumes substantial amounts of time, with possible additional time taken up by read receipts (associated with 13.8% of the e-mails in this study). Not surprisingly, the general trend was toward more e-mails for those with higher academic rank and more publications. Either factor would be associated with more visibility online including disclosure of institutional e-mail addresses.

The journals in unsolicited e-mails had a number of characteristics observed in previous studies. Two common characteristics are claims of article peer review in the e-mail body and low rates of inclusion in journal databases that have clearly defined inclusion and exclusion criteria (MEDLINE, PubMed Central, EMBASE, Scopus, JCR, CINAHL, and DOAJ).[16,19,20] Of the 328 unique journal names captured in our study, 76.4% were not found in journal lists for the databases/resources analyzed. While 17.7% of the journals were included in PubMed, 12.5% were in PubMed solely as a result of being NIH-funded research and directly deposited by the author into PubMed Central. This route has been identified as a mechanism for potentially low-quality journals to have articles included in PubMed.[18]

Similar to a detailed analysis of the characteristics of potentially predatory journals,[16] we found that promotion of Index Copernicus and its associated Index Copernicus Value was one of the more common claims in the e-mails related to journals. The validity of the Index Copernicus has been challenged;[16] although to the best of our knowledge, there has not been detailed published analysis of the journals contained in this index. Claiming inclusion in Index Copernicus is also somewhat ambiguous as there is both an Index Copernicus International (ICI) World of Journals and an ICI Journals Masters List.[21] The ICI World of Journals simply requires registration by the journal publisher and is said to contain over 45,000 scientific journals. The ICI Journals Master Lists also requires registration of the journal by the publisher and now contains over 6500 journals that have been reviewed by ICI and assigned an Index Copernicus Value, a proprietary measure used by ICI. ICI states that their index value focuses on a journal's potential for future citations rather than scoring actual citations and thus scores rise according to factors such as "cooperation," "internationalization," and "digitization."[22] The highest-ranked journal by Index Copernicus Value in 2017 was *Comunicar*, an international "educommunication" journal with an average-for-category Journal Impact Factor of 2.8 according to Clarivate Analytics. By contrast, *The New England Journal of Medicine*, with a third of a million actual citations, had a 2017 Journal Impact Factor of 79.3 (second highest of all journals) according to Clarivate Analytics yet was unlisted by Index Copernicus. In fact, the journals with the highest 20 Journal Impact Factors in 2017 in Clarivate Analytics are all not included in Index Copernicus Journal Masters List (possibly simply due to the publishers of these journals not choosing to register with this index), illustrating both the difference in ranking and selective nature of the Index Copernicus compared to a more widely accepted numeric "impact" or "index" value.

The adjective "predatory" has been applied to academic journals and conferences of potentially poor quality and scientific value.[1-7] Most of the literature to date has focused on predatory journals, with some studies aimed at identifying key characteristics of these journals.[16,19,23-25] Along these lines, one common approach is to determine if a journal is actually included in established databases/resources that have defined criteria and review practices. One limitation is that it takes time for journals to meet the criteria for inclusion. This can make the evaluation of newer journals difficult to assess in the short term.[26] The other approach is the development of "black lists" of potentially predatory or low-quality journals.

The blacklist approach was exemplified by Beall's List, a now-defunct list of potentially predatory journals assembled and published by a University of Colorado librarian.[27] Since the discontinuation of Beall's list in 2017, a group of anonymous contributors puts together a website with potentially predatory journals which we have used in the current study ["Stop Predatory Journals", Table 1].[17] There is also a commercially available blacklist that we did not have access to in our study.[28]

There are several limitations to the blacklist approach. First, the reasons why a journal appears on a blacklist may not be transparent. Related to this, there may also not be a clear way to adjudicate disputes by a journal questioning inclusion on the list. Second, the rapid growth of journals makes it very challenging logistically to keep up with the pace of new entities. There is also the possibility that quality for a journal improves or degrades when acquired by a different publisher or with changes

in editorial control.[29] Third, there is difficulty in evaluating new journals that have not yet published a sufficient number of articles to assess quality. Some of these will become established and of sound quality, but this can be hard to assess in the early phases of the journal.[26] Fourth, journals inherently vary in intent and purpose. In this regard, some journals are catering more to researchers in developing nations that may encounter barriers in publishing and/or to more speculative or controversial content difficult to publish in traditional journals.[30,31] Finally, labeling a journal or publisher as potentially predatory carries the risk of legal or other action by publishers.

Analogous to journals, concern has also arisen regarding conferences or meetings that are of potentially poor quality and scientific value.[3,12-14,30] Some of these may be outright fraudulent (including the possibility that the event does not even exist) or a meeting of low quality. There is much less systematic study of these conferences and meetings. A common challenge is that journals, conferences, and webinars have international scope, making it difficult to have global regulations or governing bodies overseeing quality.

In summary, the present study shows substantial burden of unsolicited e-mails from journals to faculty and trainees in an academic pathology department. As such, student, resident, and fellow education should include information helpful in assessing the quality of journals and conferences along with strategies for choosing journals for publishing. Limitations of the present study include a single academic institution and analysis of only a subset of trainees and faculty. The study also does not capture the full range of unsolicited e-mails that include newsletters, advertisements, and announcements. Nevertheless, this is the first study to look systematically at e-mails received by individuals at a range of academic ranks in a medical center.

Conclusion

We have demonstrated that faculty and trainees in an academic pathology department receive a high burden of unsolicited e-mails from medical/scientific journals, conferences, and webinars. The number of e-mails correlates with higher academic rank and number of publications. Less than half of the e-mails are recognized as spam by routine institutional e-mail settings.

Acknowledgments

The authors would like to thank Tony J. Wissink (University of Iowa Hospitals and Clinics Department of Pathology) for helpful advice on technical aspects of analyzing the emails.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

1. Bartholomew RE. Science for sale: The rise of predatory journals. J R Soc Med 2014;107:384-5.
2. Beall J. Medical publishing and the threat of predatory journals. Int J Womens Dermatol 2016;2:115-6.
3. Bowman JD. Predatory publishing, questionable peer review, and fraudulent conferences. Am J Pharm Educ 2014;78:176.
4. Butler D. Investigating journals: The dark side of publishing. Nature 2013;495:433-5.
5. Eriksson S, Helgesson G. The false academy: Predatory publishing in science and bioethics. Med Health Care Philos 2017;20:163-70.
6. Harvey HB, Weinstein DF. Predatory publishing: An emerging threat to the medical literature. Acad Med 2017;92:150-1.
7. Roberts J. Predatory journals: Illegitimate publishing and its threat to all readers and authors. J Sex Med 2016;13:1830-3.
8. Moher D, Shamseer L, Cobey KD, Lalu MM, Galipeau J, Avey MT, *et al.* Stop this waste of people, animals and money. Nature 2017;549:23-5.
9. Moher D, Srivastava A. You are invited to submit…. BMC Med 2015;13:180.
10. Bohannon J. Who's afraid of peer review? Science 2013;342:60-5.
11. Hawkes N. Spoof research paper is accepted by 157 journals. BMJ 2013;347:f5975.
12. Darbyshire P. Fake news. Fake journals. Fake conferences. What we can do. J Clin Nurs 2018;27:1727-9.
13. Mercier E, Tardif PA, Moore L, Le Sage N, Cameron PA. Invitations received from potential predatory publishers and fraudulent conferences: A 12-month early-career researcher experience. Postgrad Med J 2018;94:104-8.
14. Sorooshian S. Conference wolves in sheep's clothing. Sci Eng Ethics 2017;23:1805-6.
15. Spam Confidence Levels. Microsoft Exchange; Updated 01 October, 2017. Available from: https://www.docs.microsoft.com/en-us/office365/securitycompliance/spam-confidence-levels. [Last accessed on 2019 Mar 06].
16. Shamseer L, Moher D, Maduekwe O, Turner L, Barbour V, Burch R, *et al.* Potential predatory and legitimate biomedical journals: Can you tell the difference? A cross-sectional comparison. BMC Med 2017;15:28.
17. Stop Predatory Journals. Available from: https://www.predatoryjournals.com/. [Last accessed on 2019 Mar 06].
18. Manca A, Moher D, Cugusi L, Dvir Z, Deriu F. How predatory journals leak into PubMed. CMAJ 2018;190:E1042-5.
19. Cobey KD, Lalu MM, Skidmore B, Ahmadzai N, Grudniewicz A, Moher D. What is a predatory journal? A scoping review. F1000Res 2018;7:1001.
20. Bolshete P. Analysis of thirteen predatory publishers: A trap for eager-to-publish researchers. Curr Med Res Opin 2018;34:157-62.
21. ICI Journals Master List; 2019. Available from: https://www.journals.indexcopernicus.com/. [Last accessed on 2019 Mar 06].
22. ICI Journals Master List. Evaluation Methdology; 2019. Available from: https://www.indexcopernicus.com/index.php/en/parametrisation-1/journals-master-list-2/the-methodology-en. [Last accessed on 2019 Mar 06].
23. Andoohgin Shahri M, Jazi MD, Borchardt G, Dadkhah M. Detecting Hijacked journals by using classification algorithms. Sci Eng Ethics 2018;24:655-68.
24. Dadkhah M, Maliszewski T, Teixeira da Silva JA. Hijacked journals, Hijacked web-sites, journal phishing, misleading metrics, and predatory publishing: Actual and potential threats to academic integrity and publishing ethics. Forensic Sci Med Pathol 2016;12:353-62.
25. Dadkhah M, Maliszewski T, Jazi MD. Characteristics of Hijacked journals and predatory publishers: Our observations in the academic world. Trends Pharmacol Sci 2016;37:415-8.
26. Singer A. Not all young journals are predatory. West J Emerg Med 2017;18:318.
27. Strielkowski W. Predatory journals: Beall's list is missed. Nature 2017;544:416.
28. Das S, Chatterjee SS. Cabell's blacklist: A new way to tackle predatory journals. Indian J Psychol Med 2018;40:197-8.
29. Spears T. Canadian cardiac journal turns "predatory". CMAJ 2014;186:E525.
30. Memon AR, Azim ME. Predatory conferences: Addressing researchers from developing countries. J Pak Med Assoc 2018;68:1691-5.
31. Memon AR, Waqas A. Indexing by bibliographic databases of journals published in the developing world. Sci Eng Ethics 2018;24:1371-5.