BMJ Open

Examining Academic Spam Emails Among Career Development Grant Awardees: The ERASE (Elucidating and Removing Academic Spam Emails) Study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-027928
Article Type:	Research
Date Submitted by the Author:	14-Nov-2018
Complete List of Authors:	Wilkinson, Tracey; Indiana University Department of Pediatrics, Children's Health Services Research Russell, Christopher; Children's Hospital of Los Angeles; Keck School of Medicine of the University of Southern California Bennett, William; Riley Hospital for Children, Pediatric Gastroenterology Cheng, Erika; Indiana University Department of Pediatrics Carroll, Aaron; Indiana University Department of Pediatrics
Keywords:	open access, publishing, time management

Examining Academic Spam Emails Among Career Development Grant Awardees: The ERASE (Elucidating and Removing Academic Spam Emails) Study

Tracey A. Wilkinson, MD, MPH[1*], Christopher J. Russell, MD, MS,[2] William E. Bennett, Jr, MD, MS[1], Erika R. Cheng, PhD, MPA[1], Aaron E. Carroll, MD, MS[3]

[1]Assistant Professor, Department of Pediatrics, Indiana University School of Medicine, Indianapolis, IN
[2]Assistant Professor, Department of Pediatrics, Keck School of Medicine, University of Southern California, Los Angeles, CA, Division of Hospital Medicine, Children's Hospital Los Angeles, Los Angeles, CA, United States
[3]Professor, Department of Pediatrics, Indiana University School of Medicine, Indianapolis, IN

* Corresponding Author:
Tracey A. Wilkinson
410 W. 10th Street, Suite 2000
Indianapolis, IN. 46202
(317)278-0552
tracwilk@iu.edu

Manuscript Word Count: 1735

ABSTRACT

Examining Academic Spam Emails Among Career Development Grant Awardees: The ERASE (Elucidating and Removing Academic Spam Emails) Study

Objective: To investigate the scope of academic spam emails amongst career development grant awardees and factors associated with the amount of time spent addressing them.

Design: A cross-sectional survey of career development grant investigators via anonymous online survey. In addition to demographic and professional information, we asked investigators to report the number of academic spam emails received each day, how they determined whether these emails were spam, and time they spent per day addressing them. We used bivariate analysis to assess factors associated with the amount of time spent on academic spam emails.

Setting: Online survey sent on three separate occasions between Nov-December, 2016.

Participants: All National Institute of Health career development awardees funded in the 2015 fiscal year.

Main Outcome Measures: Factors associated with the amount of time spent addressing academic spam emails.

Results: A total of 3492 surveys were emailed, of which 206 (5.9%) were returned as undeliverable and 96 (2.7%) reported an out-of-office message; our overall response rate was 22.3% (n=733). All respondents reported receiving academic spam emails, with the majority (54.4%) receiving between 1 and 10 per day and spending between 1 and 10 minutes each day evaluating them. The amount of time respondents reported spending on academic spam emails was associated with the number of peer-reviewed journal articles authored ($p<0.001$), a history of publishing in open access format ($p<0.01$), the total number of academic spam emails received ($p<0.001$), and a feeling of having missed opportunities due to ignoring these emails ($p=0.04$).

Conclusions: Academic spam emails are a common distraction for career development grantees that may impact faculty productivity. There is an urgent need to mitigate this growing problem.

Keywords: Open Access Publishing, Publishing, Time Management

Strengths and Limitations of This Study

- This is the first study to describe the scope of academic spam emails amongst career development grantees
- The survey was distributed by email, thus could have been perceived as a spam email by recipients
- The survey only included recipients of National Institutes of Health funding, and did not include those that have applied for funding and were not successful

INTRODUCTION

The number of open access journals, and subsequent published articles, has increased dramatically in the last 10 years, with estimated revenues for the academic publishing industry of approximately $10.5 billion per year globally.[1-4] The original goal of open access was to enhance access to research for the general public and other researchers and has been successful in that respect. However, a subset of journals and publishers that have been deemed to be "predatory" have also been able to flourish in this environment.

These "predatory" publishers send unsolicited emails requesting manuscript submissions, offer rapid review, and use publication fees (rather than traditional peer or editorial review) to select articles. As a result of this model, there have been documented instances of accepting flawed manuscripts.[5-9] Because academic research faculty career success is often centered on their publication and funding record, their ability to obtain funding to conduct research is intimately linked to their success in publishing their work. Junior academic faculty are particularly focused on opportunities to publish given the importance tied to their promotion. They may, therefore, be most susceptible to academic spam emails (ASEs), which contain unsolicited requests for manuscripts, presentations at organization meetings, and memberships on editorial boards.

Previous studies have examined the quantity and quality of predatory open access journals, as well as characteristics of authors who publish in them.[10-16] However, to our knowledge, no study has described the prevalence of this phenomenon amongst researchers early in their career or quantified the time spent on ASEs. Therefore, we

sought to examine factors associated with the amount of time spent addressing academic spam emails by career development grantees.

METHODS

We obtained the contact information for all National Institution of Health (NIH) K-awardees in the fiscal year of 2015 via a Freedom of Information Act request. We then created and emailed an online survey using Survey Monkey[17] consisting of 14 questions to the correspondent/recipient on record of each K-award. The survey was developed by the authors and piloted amongst faculty at our institution that were not current K-awardees to assure clarity of questions and adequate response options. The survey was sent a total of 3 times over a 2 month period in 2016. The email was generated from a valid, personal email account of one of the study investigators.

The survey requested information regarding basic demographic (e.g., gender) and professional information (e.g., academic job title, year of terminal degree, and publication history). Survey respondents were also asked several questions about their experience with academic spam emails—which was described to survey respondents as an "unsolicited email requesting articles/editorial, conference presentations or editorial membership." These questions asked the number of these emails they received daily, how they determined whether these emails were spam, the amount of time spent on these emails, whether their email account had a spam filter, and if they ever felt like they missed opportunities due to ignoring these emails.

We summarized responses to these questions using descriptive statistics and then assessed factors associated with the daily amount of time spent on these emails (re-categorized as none, 1-10 minutes and greater than 10 minutes) using bivariate analysis.

We removed respondents with missing data (n=68) or who replied "not sure" (n=4) to the question about time spent daily addressing academic spam emails from the bivariate analysis. Respondents removed from this analysis were less likely to have more than 20 publications, but otherwise did not differ from those included in the sample by key study variables. The Indiana University IRB approved this study with a waiver of consent.

RESULTS

A total of 3492 surveys were emailed, of which 206 (5.9%) were immediately returned as undeliverable; 96 (2.7%) of these 3286 emails had an out-of-office message automatically in response to the inquiry. Of the emails sent and received successfully, 733 (22.3%) surveys were completed by the end of the study period. A slight majority were female (n=399; 54.4%). Over two-thirds (65.5%) of respondents were assistant professors working at an academic center; most (71.9%) received their NIH-K award funding in the 4 years preceding the survey (Table 1).

Table 1. Description of Survey Respondents

Variable	Total N (%)
Total Reponses	733 (100)
Demographics and Professional Information	
Gender	
Female	399 (54.4)
Male	322 (43.9)
Other/Prefer Not to Answer	12 (1.6)
Academic Job Title (n=731)	
Assistant Professor	479 (65.5)
Associate Professor	120 (16.4)
Professor	49 (6.7)
Non Tenure-Track	57 (7.8)
Other	26 (3.6)
First Year to Receive NIH Funding (n=729)	
2011-2016	524 (71.9)
2010-2000	169 (23.3)

Prior to 2000	36 (4.9)
Number of Academic Publications (n=732)	
1 to 10	55 (7.5)
11 to 20	181 (24.7)
Greater than 20	496 (67.8)
Previously Published in Open Access Format (n=730)	
Yes	489 (67.0)
No	195 (26.7)
Not Sure	46 (6.3)
Year Terminal Degree Obtained (n=729)	
2010-2016	137 (18.8)
2005-2009	307 (42.1)
2000-2004	202 (27.7)
Prior to 2000	83 (11.4)
SPAM Filter on Email (n=729)	
Yes	651 (89.3)
No	36 (4.9)
Not Sure	42 (5.8)
Experience with Academic Spam Emails	
Estimated Number of Spam Emails/Day (n=724)	
1 to 10	398 (54.5)

Table 1. Description of Survey Respondents (continued)

Variable	**Total N (%)**
11 to 20	223 (30.6)
Greater than 20	103 (14.1)
Not Sure	6 (0.8)
Time Spent Per Day Reading/Sorting Spam Emails (n=665)	
None, delete them all without reading	132 (19.9)
1 to 10 minutes	419 (63.0)
11 to 20 minutes	89 (13.4)
More than 20 minutes	21 (3.2)
Not Sure	4 (0.6)
History of Responding to Academic Spam Email (n=729)	
Yes	226 (31.0)
No	503 (69.0)

Concern of Missed Opportunities (n=731)	
Yes	58 (7.9)
No	575 (78.7)
Not Sure	98 (13.4)
Methods used to Deem Emails Academic Spam[a]	
Consider them all Spam	274 (37.4)
Ask a Colleague	123 (16.8)
Typos in text/name	398 (54.5)
Don't Recognize Journal	601 (82.0)
Address listed is not in US	308 (42.0)
Requesting Fee to Publish	313 (42.7)
Consult the Internet	239 (32.6)

a. Respondents were instructed to select all that apply.
Note. Percentages may not sum to 100 due to rounding

With respect to number of ASEs received each day, every respondent reported receiving at least one academic spam email in the previous week (data not shown). Over half (54.5%) reported receiving between 1-10 academic spam emails per day, 30.6% reported receiving between 11 and 20 emails per day, and 14.1% reported receiving more than 20 emails per day (Table 1). When asked how much time was spent on these emails in a day, 80% of the respondents reported spending at least some amount of time during the day addressing these emails, with 63% of respondents reported spending between 1 and 10 minutes of their day reading, sorting, and determining what to do with potential ASEs. The methods of how they determined an email to be academic spam were diverse, with an unknown journal name being the most frequently described method (82% of respondents). More than a one third of respondents (31%) reported that they had responded to an ASE in the past.

On bivariate analysis, when examining variables associated with the amount of time spent on academic spam emails in a day (Table 2), we found that neither gender (overall p-value=0.37) nor academic job title (overall p-value=0.36) were associated with

the time spent addressing academic spam emails. Interestingly, the time spent on emails was also not associated with reporting having a SPAM filter (p=0.99).

Table 2. Factors associated with time spent on emails

	Time Spent on Emails Per Day (%)			
	None	1-10min	>10min	chi-square p-value
Demographics and Professional Information				
Gender				0.37
Female	19.9	66.0	14.0	
Male	20.1	60.4	19.5	
Other/Prefer not to answer	16.7	58.3	25.0	
Academic Job Title				
Assistant Professor	19.5	65.4	15.1	0.36
Associate Professor	20.7	57.8	21.6	
Professor	28.6	51.0	20.4	
Non-Tenure Track	17.4	65.2	17.4	
Other	10.5	79.0	10.5	
Number of Academic Publications				
1 to 10	0.0	87.5	12.5	<0.001
11 to 20	0.0	80.8	19.2	
Greater than 20	26.9	56.8	16.3	
Previously Published in Open Access Format				
Yes	23.7	67.5	8.9	0.01
No	19.7	61.7	18.6	
Not Sure	9.3	65.1	25.6	
SPAM Filter on Email				
No	19.9	63.1	17.0	0.99
Yes	20.7	65.5	13.8	
Not Sure	20.0	65.0	15.0	
Experience with Academic Spam Emails				
Estimated Number of Spam Emails/Day				
1 to 10	20.3	72.9	6.8	<0.001
11 to 20	20.3	58.0	21.7	
Greater than 20	18.6	40.2	41.2	
Not Sure	0.0	100.0	0.0	
Concern for Missed Opportunities				

No	21.3	63.8	14.9	0.04
Yes	14.3	55.4	30.4	
Not Sure	15.9	65.9	18.2	

However, there was an inverse association (overall p-value <0.001) with the number of academic publications reported and the time spent addressing these emails, such that faculty with fewer publications reported spending more time per day reading potential ASEs and assessing their legitimacy. Among faculty with 10 or fewer publications, 87.5% spent between 1 and 10 minutes assessing these emails compared with those with 11-20 publications (80.8%) and greater than 20 (56.8%) publications, respectively. More than a quarter of respondents with greater than 20 publications spent no time assessing these emails compared with 0% of respondents with fewer publications. There was also an association between the number of potential ASEs received in a day and the time spent on them (overall p-value <.001). Specifically, respondents who received more than 20 ASEs per day were more likely to report spending more than 10 minutes per day on these emails than respondents who received fewer emails.

Finally, we assessed whether respondents felt they might have missed opportunities in the past due because they mistakenly assumed legitimate emails were academic spam. We found that respondents who were concerned about missed opportunities were almost twice as likely to spend at least 10 minutes of their day assessing these emails than those who were not concerned (24.3% versus 12.8%; overall p-value=0.04).

DISCUSSION

This study sought to examine the scope of ASEs among NIH career development awardees and factors associated with the amount of daily time spent addressing them. We found that receiving ASEs was pervasive in this population. In fact, everyone who responded to the survey reported receiving academic spam, precluding us from examining demographic and job-related factors associated with the receipt of these emails. In addition, 80% of respondents reported using time during their day to address these emails.

We did examine these factors in relation to the amount of time respondents spent addressing the emails in any given day. Neither gender, academic rank, nor having a spam filter affected the time spent addressing these emails, but faculty with less than 20 publications and who felt like they might be missing opportunities did report spending more time addressing these emails than their counterparts. More than 30% of respondents who felt that they might have missed opportunities for publication, presentation, or editorial service because they ignored an academic spam email reported spending more than 10 minutes every day reading and sorting academic spam. This suggests that academic faculty who are junior and/or who feel pressure to publish might be most susceptible to these types of predatory solicitations. In addition, the more ASEs received in a day was associated with more time spent addressing them.

Our study has some limitations. First, we used the contact information for grant awardees provided by the NIH and did not determine individual contact information based on awardees name and institution separately. Thus, some email addresses were not working or valid at the time of our survey and may have represented an institutional contact and not the awardee themselves. Furthermore, our study sample was those with

funded NIH grants, not including those that may have applied but were not funded. Second, there is a chance that our emailed survey was viewed by recipients as SPAM itself, and possibly ignored or deleted without a response. We used a valid personal email address with a survey link in the body of the email to solicit respondents, but that also prohibited us from knowing who had completed the survey already or who to follow-up with directly for not completing the survey. Third, the overall response rate was 22% and we did not require any question on the survey to be answered and thus had some missing data. However, respondents removed from the sample for missing data did not systematically differ from those included, except that they reported having fewer publications. This likely produced conservative results in our bivariate analysis.

Nonetheless, our study quantifies the burden of time faculty spend addressing ASEs as career development grantees. As the number of journals and publishers continue to rise, we can expect the number of emails to do so as well. Given the focus on faculty to publish, it is likely that they will continue to spend time reading these emails to determine whether they are legitimate or not, when they could be working on projects, manuscripts, and grant proposals. Knowing that ASEs will likely not stop, this may indicate an additional piece of training faculty should receive and have resources given to determine whether or not to respond to unsolicited emails or any easy way to report them to help identify the senders as spam in the future. There are efforts underway to help researchers choose the right journal for their research which includes deciphering predatory journals from others.[18] Furthermore, efforts at the university level to help improve email filters to block unsolicited emails can be refined and improved. From a

larger system perspective, the publishing industry can build upon their current communications to help differentiate from predatory journals for the receivers.

Acknowledgements: We would like to thank all the investigators that responded to this survey and Professor Jeffrey Beall who helped with the concept of this study. None of the authors have conflicts of interest or financial disclosures to report.

Competing Interests: None

Funding: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors

Author Contributions:
Dr. Wilkinson conceptualized and designed the study, distributed the survey, drafted the initial manuscript and approved the final manuscript as submitted.

Dr. Russell conceptualized and designed the study and approved the final manuscript as submitted

Dr. Bennett conceptualized and designed the study and approved the final manuscript as submitted

Dr. Cheng conceptualized and designed the study, carried out the analysis and approved the final manuscript as submitted.

Dr. Carroll contributed to the study design and data collection, reviewed and revised the manuscript, and approved of the final manuscript as submitted.

Data Sharing Statement: The survey, statistical code and dataset available by request to corresponding author, Dr. Tracey Wilkinson

References

1. Esposito P. The size of the open access market. The Scholarly Kitchen; 2014. http://scholarlykitchen.sspnet.org/2014/10/29/the-size-of-the-open-access-market.
2. Shen C, Björk B-C. 'Predatory' open access: a longitudinal study of article volumes and market characteristics. BMC Med. 2015;13(1):230.
3. Eriksson S, Helgesson G. The false academy: predatory publishing in science and bioethics. Med Health Care Philos. 2017;20(2):163-70.
4. Gasparyan AY, Yessirkepov M, Diyanova SN, Kitas GD. Publishing Ethics and Predatory Practices: A Dilemma for All Stakeholders of Science Communication. J Korean Med Sci. 2015;30(8):1010-6.
5. Beall J. Criteria for determining predatory open-access publishers. 2nd edition. Denver, CO: Scholarly Open Access; 2012. http://scholarlyoa.files.wordpress.com/2012/11/criteria-2012-2.pdf.
6. Daivs P. Open Access Publisher Accepts Nonsense Manuscript for Dollars 2009 [Available from: https://scholarlykitchen.sspnet.org/2009/06/10/nonsense-for-dollars/.
7. Bohannon J. Who's afraid of peer review. Science. 2013;342(6154).
8. Beall J. Predatory publishers are corrupting open access. Nature. 2012;489(7415):179.
9. Butler D. Investigating journals: The dark side of publishing. Nature. 2013;495(7442):433-5.
10. Nwagwu WE, Ojemeni O. Penetration of Nigerian predatory biomedical open access journals 2007–2012: a bibliometric study. Learned Publishing. 2015;28(1):23-34.
11. Xia J. Predatory journals and their article publishing charges. Learned Publishing. 2015;28(1):69-74.
12. Xia J, Harmon JL, Connolly KG, Donnelly RM, Anderson MR, Howard HA. Who publishes in "predatory" journals? Journal of the Association for Information Science and Technology. 2015;66(7):1406-17.
13. Tin L, Ivana B, Biljana B, Ljubica IB, Dragan M, Dušan S. Predatory and fake scientific journals/publishers–a global outbreak with rising trend: a review. Geographica Pannonica. 2014;18(3):69-81.
14. Sorokowski P, Kulczycki E, Sorokowska A, Pisanski K. Predatory journals recruit fake editor. Nature. 2017;543(7646):481-3.
15. Cobey K. Illegitimate journals scam even senior scientists. Nature. 2017;549(7670):7.
16. Bolshete P. Analysis of thirteen predatory publishers: a trap for eager-to-publish researchers. Curr Med Res Opin. 2018;34(1):157-62.
17. SurveyMonkey, Inc. San Mateo, California, USA.
18. Beall J. Best practices for scholarly authors in the age of predatory journals. Ann R Coll Surg Engl. 2016;98(2):77-9.

BMJ Open

A Cross-Sectional Study of Predatory Publishing Emails Received by Career Development Grant Awardees: The ERASE Study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-027928.R1
Article Type:	Research
Date Submitted by the Author:	16-Jan-2019
Complete List of Authors:	Wilkinson, Tracey; Indiana University Department of Pediatrics, Children's Health Services Research Russell, Christopher; Children's Hospital of Los Angeles; Keck School of Medicine of the University of Southern California Bennett, William; Riley Hospital for Children, Pediatric Gastroenterology Cheng, Erika; Indiana University Department of Pediatrics Carroll, Aaron; Indiana University Department of Pediatrics
Primary Subject Heading:	Medical publishing and peer review
Secondary Subject Heading:	Medical education and training
Keywords:	open access, publishing, time management, predatory journals, open access publishing

A Cross-Sectional Study of Predatory Publishing Emails Received by Career Development Grant Awardees: The ERASE Study

Tracey A. Wilkinson, MD, MPH[1*], Christopher J. Russell, MD, MS,[2] William E. Bennett, Jr, MD, MS[1], Erika R. Cheng, PhD, MPA[1], Aaron E. Carroll, MD, MS[3]

[1]Assistant Professor, Department of Pediatrics, Indiana University School of Medicine, Indianapolis, IN
[2]Assistant Professor, Department of Pediatrics, Keck School of Medicine, University of Southern California, Los Angeles, CA, Division of Hospital Medicine, Children's Hospital Los Angeles, Los Angeles, CA, United States
[3]Professor, Department of Pediatrics, Indiana University School of Medicine, Indianapolis, IN

* Corresponding Author:
Tracey A. Wilkinson
410 W. 10th Street, Suite 2000
Indianapolis, IN. 46202
(317)278-0552
tracwilk@iu.edu

Manuscript Word Count: 2012

ABSTRACT

A Cross-Sectional Study of Predatory Publishing Emails Received by Career Development Grant Awardees: The ERASE Study

Objective: To investigate the scope of academic spam emails amongst career development grant awardees and factors associated with the amount of time spent addressing them.

Design: A cross-sectional survey of career development grant investigators via anonymous online survey was conducted. In addition to demographic and professional information, we asked investigators to report the number of academic spam emails received each day, how they determined whether these emails were spam, and time they spent per day addressing them. We used bivariate analysis to assess factors associated with the amount of time spent on academic spam emails.

Setting: An online survey sent via email on three separate occasions between Nov-December, 2016.

Participants: All National Institute of Health career development awardees funded in the 2015 fiscal year.

Main Outcome Measures: Factors associated with the amount of time spent addressing academic spam emails.

Results: A total of 3492 surveys were emailed, of which 206 (5.9%) were returned as undeliverable and 96 (2.7%) reported an out-of-office message; our overall response rate was 22.3% (n=733). All respondents reported receiving academic spam emails, with the majority (54.4%) receiving between 1 and 10 per day and spending between 1 and 10 minutes each day evaluating them. The amount of time respondents reported spending on academic spam emails was associated with the number of peer-reviewed journal articles authored ($p<0.001$), a history of publishing in open access format ($p<0.01$), the total number of academic spam emails received ($p<0.001$), and a feeling of having missed opportunities due to ignoring these emails ($p=0.04$).

Conclusions: Academic spam emails are a common distraction for career development grantees that may impact faculty productivity. There is an urgent need to mitigate this growing problem.

Keywords: Open Access Publishing, Predatory Journals Publishing, Time Management

Strengths and Limitations of This Study

- This is the first study to describe the scope of academic spam emails amongst career development grantees
- The survey was distributed by email, thus could have been perceived as a spam email by recipients
- The survey only included recipients of National Institutes of Health funding, and did not include those that have applied for funding and were not successful

INTRODUCTION

The number of open access journals, and subsequent published articles, has increased dramatically in the last 10 years, with estimated revenues for the academic publishing industry of approximately $10.5 billion per year globally.[1-4] The original goal of open access was to enhance access to research for the general public and other researchers and has been successful in that respect. However, a subset of journals and publishers that have been deemed to be "predatory" have also been able to flourish in this environment.

This subset of "predatory" publishers send unsolicited emails that request manuscript submissions, offer rapid review, and use publication fees (rather than traditional peer or editorial review) as criteria to accept articles.[5-7] As a result of this model, there have been documented instances of accepting flawed manuscripts or fake editorial board members.[8-11] Because academic research faculty career success is often centered on their publication and funding record, their ability to obtain funding to conduct research is intimately linked to their success in publishing their work. Junior academic faculty are particularly focused on opportunities to publish given the importance tied to their promotion. They may, therefore, be most susceptible to academic spam emails (ASEs) from predatory publishers, which contain unsolicited requests for publishing manuscripts, presentations at organization meetings, and memberships on editorial boards.

Previous studies have examined the quantity and quality of predatory publishers, as well as characteristics of authors who publish in them.[10, 12-19] However, to our knowledge, no study has described the prevalence of this phenomenon amongst

researchers early in their career or quantified the time spent on ASEs. Therefore, we sought to examine factors associated with the amount of time spent addressing academic spam emails by career development grantees in the ERASE (Elucidating and Removing Academic Spam Emails) Study.

METHODS

Study Design and Data Source

We obtained the contact information for all National Institution of Health (NIH) K-awardees in the fiscal year of 2015 via a Freedom of Information Act request. Information of federal grant awardees is a standard report issued by the NIH Freedom of Information Office that is available online.[20] We then created and emailed an online survey using Survey Monkey[21] consisting of 14 questions to the correspondent/recipient on record of each K-award.

The survey was developed by the authors and piloted amongst 10 faculty at Indiana University School of Medicine that were not current K-awardees to assure clarity of questions and adequate response options. If a question was not clear or the answer choices not sufficient, the survey was edited before it was sent to the next faculty member. This process was repeated until there were no additional feedback suggestions. The survey was sent a total of 3 times over a 2 month period (November-December) in 2016. The email was generated from a valid, personal email account of one of the study investigators (TAW).

The survey requested information regarding basic demographic (e.g., gender) and professional information (e.g., academic job title, year of terminal degree, and publication history). Survey respondents were also asked several questions about their experience

with academic spam emails—which was described to survey respondents as an "unsolicited email requesting articles/editorial, conference presentations or editorial membership." Subsequent questions asked respondents to report the number of academic spam emails they received daily, how they determined whether these emails were spam, the amount of time spent on these emails, whether their email account had a spam filter, and if they ever felt like they missed opportunities due to ignoring these emails (supplementary file).

Data Analysis

We summarized responses to survey questions using descriptive statistics and then assessed factors associated with the daily amount of time spent on these emails (re-categorized as none, 1-10 minutes and greater than 10 minutes) using chi-squared bivariate analysis. We removed respondents with missing data (n=68) or who replied "not sure" (n=4) to the question about time spent daily addressing academic spam emails from the bivariate analysis. Respondents removed from this analysis were less likely to have more than 20 publications, but otherwise did not differ from those included in the sample by key study variables. All analysis was done use SAS Version 9.4 (SAS Institute INC, Cary, NC).

Ethical Approval

The Indiana University IRB approved this study with a waiver of consent.

Patient and Public Involvement

For this study, there was no patient involvement; however, we collected data through a survey of NIH grantees available from a public database.[20]

RESULTS

A total of 3492 surveys were emailed, of which 206 (5.9%) were immediately returned as undeliverable; 96 (2.7%) of these 3286 emails had an out-of-office message automatically in response to the inquiry. Of the emails sent and received successfully, 733 (22.3%) surveys were completed by the end of the study period. A slight majority were female (n=399; 54.4%). Over two-thirds (65.5%) of respondents were assistant professors working at an academic center; most (71.9%) received their NIH-K award funding in the 4 years preceding the survey (Table 1).

Table 1. Description of Survey Respondents*

Variable	Total N (%)
Total Reponses	733 (100)
Demographics and Professional Information	
Gender	
Female	399 (54.4)
Male	322 (43.9)
Other/Prefer Not to Answer	12 (1.6)
Academic Job Title (n=731)	
Assistant Professor	479 (65.5)
Associate Professor	120 (16.4)
Professor	49 (6.7)
Non Tenure-Track	57 (7.8)
Other	26 (3.6)
First Year to Receive NIH Funding (n=729)	
2011-2016	524 (71.9)
2010-2000	169 (23.3)
Prior to 2000	36 (4.9)
Number of Academic Publications (n=732)	
1 to 10	55 (7.5)
11 to 20	181 (24.7)
Greater than 20	496 (67.8)
Previously Published in Open Access Format (n=730)	
Yes	489 (67.0)
No	195 (26.7)
Not Sure	46 (6.3)
Year Terminal Degree Obtained (n=729)	

2010-2016	137 (18.8)
2005-2009	307 (42.1)
2000-2004	202 (27.7)
Prior to 2000	83 (11.4)
SPAM Filter on Email (n=729)	
Yes	651 (89.3)
No	36 (4.9)
Not Sure	42 (5.8)
Experience with Academic Spam Emails	
Estimated Number of Spam Emails/Day (n=724)	
1 to 10	398 (54.5)

Table 1. Description of Survey Respondents (continued)

Variable	**Total N (%)**
11 to 20	223 (30.6)
Greater than 20	103 (14.1)
Not Sure	6 (0.8)
Time Spent Per Day Reading/Sorting Spam Emails (n=665)	
None, delete them all without reading	132 (19.9)
1 to 10 minutes	419 (63.0)
11 to 20 minutes	89 (13.4)
More than 20 minutes	21 (3.2)
Not Sure	4 (0.6)
History of Responding to Academic Spam Email (n=729)	
Yes	226 (31.0)
No	503 (69.0)
Concern of Missed Opportunities (n=731)	
Yes	58 (7.9)
No	575 (78.7)
Not Sure	98 (13.4)
Methods used to Deem Emails Academic Spam[a]	
Consider them all Spam	274 (37.4)
Ask a Colleague	123 (16.8)
Typos in text/name	398 (54.5)
Don't Recognize Journal	601 (82.0)
Address listed is not in US	308 (42.0)

Requesting Fee to Publish	313 (42.7)
Consult the Internet	239 (32.6)

a. Respondents were instructed to select all that apply.
Note. Percentages may not sum to 100 due to rounding, total n differs if a question was left blank

With respect to number of ASEs received each day, every respondent reported receiving at least one academic spam email in the previous week (data not shown). Over half (54.5%) reported receiving between 1-10 academic spam emails per day, 30.6% reported receiving between 11 and 20 emails per day, and 14.1% reported receiving more than 20 emails per day (Table 1). When asked how much time was spent on these emails in a day, 80% of the respondents reported spending at least some amount of time during the day addressing these emails, with 63% of respondents reported spending between 1 and 10 minutes of their day reading, sorting, and determining what to do with potential ASEs. The methods of how they determined an email to be academic spam were diverse, with an unknown journal name being the most frequently described method (82% of respondents). More than a one third of respondents (31%) reported that they had responded to an ASE in the past.

On bivariate analysis, when examining variables associated with the amount of time spent on academic spam emails in a day (Table 2), we found that neither gender (p-value=0.37) nor academic job title (p-value=0.36) were associated with the time spent addressing academic spam emails. Interestingly, the time spent on emails was also not associated with reporting having a SPAM filter (p=0.99).

Table 2. Factors associated with time spent on emails

	Time Spent on Emails Per Day (%)			
	None	1-10min	>10min	p-value
Demographics and Professional Information				

Gender				0.37
Female	19.9	66.0	14.0	
Male	20.1	60.4	19.5	
Other/Prefer not to answer	16.7	58.3	25.0	
Academic Job Title				
Assistant Professor	19.5	65.4	15.1	0.36
Associate Professor	20.7	57.8	21.6	
Professor	28.6	51.0	20.4	
Non-Tenure Track	17.4	65.2	17.4	
Other	10.5	79.0	10.5	
Number of Academic Publications				
1 to 10	0.0	87.5	12.5	<0.001
11 to 20	0.0	80.8	19.2	
Greater than 20	26.9	56.8	16.3	
Previously Published in Open Access Format				
Yes	23.7	67.5	8.9	0.01
No	19.7	61.7	18.6	
Not Sure	9.3	65.1	25.6	
SPAM Filter on Email				
No	19.9	63.1	17.0	0.99
Yes	20.7	65.5	13.8	
Not Sure	20.0	65.0	15.0	
Experience with Academic Spam Emails				
Estimated Number of Spam Emails/Day				
1 to 10	20.3	72.9	6.8	<0.001
11 to 20	20.3	58.0	21.7	
Greater than 20	18.6	40.2	41.2	
Not Sure	0.0	100.0	0.0	
Concern for Missed Opportunities				
No	21.3	63.8	14.9	0.04
Yes	14.3	55.4	30.4	
Not Sure	15.9	65.9	18.2	

However, there was an inverse association (p-value <0.001) with the number of academic publications reported and the time spent addressing these emails, such that faculty with fewer publications reported spending more time per day reading potential ASEs and assessing their legitimacy. Among faculty with 10 or fewer publications,

87.5% spent between 1 and 10 minutes assessing these emails compared with those with 11-20 publications (80.8%) and greater than 20 (56.8%) publications, respectively. More than a quarter of respondents with greater than 20 publications spent no time assessing these emails compared with 0% of respondents with fewer publications. There was also an association between the number of potential ASEs received in a day and the time spent on them (p-value <.001). Specifically, respondents who received more than 20 ASEs per day were more likely to report spending more than 10 minutes per day on these emails than respondents who received fewer emails.

Finally, we assessed whether respondents felt they might have missed opportunities in the past due because they mistakenly assumed legitimate emails were academic spam. We found that respondents who were concerned about missed opportunities were almost twice as likely to spend at least 10 minutes of their day assessing these emails than those who were not concerned (24.3% versus 12.8%; p-value=0.04).

DISCUSSION

This study sought to examine the scope of ASEs among NIH career development awardees and factors associated with the amount of daily time spent addressing them. We found that receiving ASEs was pervasive in this population. In fact, everyone who responded to the survey reported receiving academic spam, precluding us from examining demographic and job-related factors associated with the receipt of these emails. In addition, 80% of respondents reported using time during their day to address these emails.

We did examine these factors in relation to the amount of time respondents spent addressing the emails in any given day. Neither gender, academic rank, nor having a spam filter affected the time spent addressing these emails, but faculty with less than 20 publications and who felt like they might be missing opportunities did report spending more time addressing these emails than their counterparts. More than 30% of respondents who felt that they might have missed opportunities for publication, presentation, or editorial service because they ignored an academic spam email reported spending more than 10 minutes every day reading and sorting academic spam. This suggests that academic faculty who are junior and/or who feel pressure to publish might be most susceptible to these types of predatory solicitations. In addition, the more ASEs received in a day was associated with more time spent addressing them.

While our study was focused on career development grantees that are generally in the early stages of their career, the burden of ASE's are not reserved for just these faculty. A 2017 study found that even senior faculty are affected by academic spam emails, making the case that efforts to control this phenomenon are needed throughout academia.[16] In addition, further work is exploring how pervasive publication within these journals is ongoing and the larger question of how academia will perceive and evaluate these publications is yet to be answered.[22, 23]

Our study has some limitations. First, we used the contact information for grant awardees provided by the NIH and did not determine individual contact information based on awardees name and institution separately. Thus, some email addresses were not working or valid at the time of our survey and may have represented an institutional contact and not the awardee themselves. Furthermore, our study sample was those with

funded NIH grants, not including those that may have applied but were not funded. Second, there is a chance that our emailed survey was viewed by recipients as SPAM itself, and possibly ignored or deleted without a response. We used a valid personal email address with a survey link in the body of the email to solicit respondents, but that also prohibited us from knowing who had completed the survey already or who to follow-up with directly for not completing the survey. Third, the overall response rate was 22% and we did not require any question on the survey to be answered and thus had some missing data. However, respondents removed from the sample for missing data did not systematically differ from those included, except that they reported having fewer publications. This likely produced conservative results in our bivariate analysis.

Nonetheless, our study quantifies the burden of time faculty spend addressing ASEs as career development grantees. As the number of journals and publishers continue to rise, we can expect the number of emails to do so as well. Given the focus of faculty to publish, it is likely that they will continue to spend time reading these emails to determine whether they are legitimate or not, when they could be working on projects, manuscripts, and grant proposals. Knowing that ASEs will likely not stop, this may indicate an additional piece of training faculty should receive and have resources given to determine whether or not to respond to unsolicited emails or any easy way to report them to help identify the senders as spam in the future.

There are efforts underway to help researchers choose the right journal for their research which includes deciphering predatory journals from others.[18, 23-27] Furthermore, efforts at the university level to help improve email filters to block unsolicited emails can be refined and improved, given a majority of our respondents had

SPAM filters in place that permitted these emails to be delivered. This may present challenges as to how to discern ASE's from other emails; however, certain patterns to emails or sender addresses could be used to flag these within a larger system. From a larger system perspective, the publishing industry can build upon their current communications to help differentiate from predatory journals for the receivers.

Acknowledgements: We would like to thank all the investigators that responded to this survey and Academic Librarian Jeffrey Beall who helped with the concept of this study. None of the authors have conflicts of interest or financial disclosures to report.

Competing Interests: None

Funding: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors

Author Contributions:
Dr. Wilkinson conceptualized and designed the study, distributed the survey, drafted the initial manuscript and approved the final manuscript as submitted.

Dr. Russell conceptualized and designed the study and approved the final manuscript as submitted

Dr. Bennett conceptualized and designed the study and approved the final manuscript as submitted

Dr. Cheng conceptualized and designed the study, carried out the analysis and approved the final manuscript as submitted.

Dr. Carroll contributed to the study design and data collection, reviewed and revised the manuscript, and approved of the final manuscript as submitted.

Data Sharing Statement: The survey, statistical code and dataset available by request to corresponding author, Dr. Tracey Wilkinson

References

1. Esposito P. The size of the open access market. The Scholarly Kitchen; 2014. http://scholarlykitchen.sspnet.org/2014/10/29/the-size-of-the-open-access-market.
2. Shen C, Björk B-C. 'Predatory' open access: a longitudinal study of article volumes and market characteristics. BMC Med. 2015;13(1):230.
3. Eriksson S, Helgesson G. The false academy: predatory publishing in science and bioethics. Med Health Care Philos. 2017;20(2):163-70.
4. Gasparyan AY, Yessirkepov M, Diyanova SN, Kitas GD. Publishing Ethics and Predatory Practices: A Dilemma for All Stakeholders of Science Communication. J Korean Med Sci. 2015;30(8):1010-6.
5. Cobey KD, Lalu MM, Skidmore B, Ahmadzai N, Grudniewicz A, Moher D. What is a predatory journal? A scoping review. F1000Research. 2018;7:1001-.
6. Beall J. Criteria for determining predatory open-access publishers. 2nd edition. Denver, CO: Scholarly Open Access; 2012. http://scholarlyoa.files.wordpress.com/2012/11/criteria-2012-2.pdf.
7. Beall J. Predatory publishers are corrupting open access. Nature. 2012;489(7415):179.
8. Daivs P. Open Access Publisher Accepts Nonsense Manuscript for Dollars 2009 [Available from: https://scholarlykitchen.sspnet.org/2009/06/10/nonsense-for-dollars/.
9. Butler D. Investigating journals: The dark side of publishing. Nature. 2013;495(7442):433-5.
10. Sorokowski P, Kulczycki E, Sorokowska A, Pisanski K. Predatory journals recruit fake editor. Nature. 2017;543(7646):481-3.
11. Bohannon J. Who's afraid of peer review. Science. 2013;342(6154).
12. Nwagwu WE, Ojemeni O. Penetration of Nigerian predatory biomedical open access journals 2007–2012: a bibliometric study. Learned Publishing. 2015;28(1):23-34.
13. Xia J. Predatory journals and their article publishing charges. Learned Publishing. 2015;28(1):69-74.
14. Xia J, Harmon JL, Connolly KG, Donnelly RM, Anderson MR, Howard HA. Who publishes in "predatory" journals? Journal of the Association for Information Science and Technology. 2015;66(7):1406-17.
15. Tin L, Ivana B, Biljana B, Ljubica IB, Dragan M, Dušan S. Predatory and fake scientific journals/publishers–a global outbreak with rising trend: a review. Geographica Pannonica. 2014;18(3):69-81.
16. Cobey K. Illegitimate journals scam even senior scientists. Nature. 2017;549(7670):7.
17. Bolshete P. Analysis of thirteen predatory publishers: a trap for eager-to-publish researchers. Curr Med Res Opin. 2018;34(1):157-62.
18. Memon AR. Predatory Journals Spamming for Publications: What Should Researchers Do? Sci Eng Ethics. 2018;24(5):1617-39.
19. 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(1):28.

20. Contact Information NIH-Supported PIs. National Institutes of Health-Freedom of Information Act Office 2015.
21. SurveyMonkey, Inc. San Mateo, California, USA.
22. Pyne D. The Rewards of Predatory Publications at a Small Business School2017. 137-60 p.
23. Eriksson S, Helgesson G. Time to stop talking about 'predatory journals'. Learned Publishing. 2018;31(2):181-3.
24. Beall J. Best practices for scholarly authors in the age of predatory journals. Ann R Coll Surg Engl. 2016;98(2):77-9.
25. Laine C, Winker MA. Identifying predatory or pseudo-journals. Biochemia medica. 2017;27(2):285-91.
26. Memon AR. Research publications and education in Pakistani medical universities: Avoiding predatory journals and improving the quality of research. J Pak Med Assoc. 2017;67(6):830-3.
27. Teixeira da Silva JA, Tsigaris P. What Value Do Journal Whitelists and Blacklists Have in Academia? The Journal of Academic Librarianship. 2018;44(6):781-92.

STROBE Checklist

Manuscript Title "A Cross-Sectional Study of Predatory Publishing Emails Received by Career Development Grant Awardees: The ERASE Study

	Item No	Recommendation	Page/Line Number
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	Title Page
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	p. 2 Abstract
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	p.3-4
Objectives	3	State specific objectives, including any prespecified hypotheses	p.4, Line 5
Methods			
Study design	4	Present key elements of study design early in the paper	p.4, Methods
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	p. 4-5, Methods
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants	p. 4-5, Methods
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	p. 4-5, Methods
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	p. 4-5, Methods
Bias	9	Describe any efforts to address potential sources of bias	p. 12, Line 3
Study size	10	Explain how the study size was arrived at	p. 4-5, Methods
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	p. 4-5, Methods
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	p. 4-5, Methods
		(*b*) Describe any methods used to examine subgroups and interactions	N/A
		(*c*) Explain how missing data were addressed	p. 5, Data Analysis
		(*d*) If applicable, describe analytical methods taking account of sampling strategy	N/A
		(*e*) Describe any sensitivity analyses	N/A
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	p. 6, Results
		(b) Give reasons for non-participation at each stage	N/A
		(c) Consider use of a flow diagram	N/A
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential	p. 6 Results, Table 1

		confounders	
		(b) Indicate number of participants with missing data for each variable of interest	Table 1
Outcome data	15*	Report numbers of outcome events or summary measures	Table 1
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	N/A
		(*b*) Report category boundaries when continuous variables were categorized	Table 1
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	N/A
Discussion			
Key results	18	Summarise key results with reference to study objectives	p. 11, Discussion
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	p. 12, Line 3
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	p. 12-13
Generalisability	21	Discuss the generalisability (external validity) of the study results	p. 12-13
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	p. 13

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

A Cross-Sectional Study of Predatory Publishing Emails Received by Career Development Grant Awardees

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-027928.R2
Article Type:	Research
Date Submitted by the Author:	13-Mar-2019
Complete List of Authors:	Wilkinson, Tracey; Indiana University Department of Pediatrics, Children's Health Services Research Russell, Christopher; Children's Hospital of Los Angeles; Keck School of Medicine of the University of Southern California Bennett, William; Indiana University School of Medicine, Pediatrics, Division of Gastroenterology Cheng, Erika; Indiana University Department of Pediatrics, Children's Health Services Research Carroll, Aaron; Indiana University Department of Pediatrics
Primary Subject Heading:	Medical publishing and peer review
Secondary Subject Heading:	Medical education and training
Keywords:	publishing, time management, predatory journals, open access publishing

A Cross-Sectional Study of Predatory Publishing Emails Received by Career Development Grant Awardees

Tracey A. Wilkinson, MD, MPH[1*], Christopher J. Russell, MD, MS,[2] William E. Bennett, Jr, MD, MS[1], Erika R. Cheng, PhD, MPA[1], Aaron E. Carroll, MD, MS[3]

[1]Assistant Professor, Department of Pediatrics, Indiana University School of Medicine, Indianapolis, IN
[2]Assistant Professor, Department of Pediatrics, Keck School of Medicine, University of Southern California, Los Angeles, CA, Division of Hospital Medicine, Children's Hospital Los Angeles, Los Angeles, CA, United States
[3]Professor, Department of Pediatrics, Indiana University School of Medicine, Indianapolis, IN

* Corresponding Author:
Tracey A. Wilkinson
410 W. 10th Street, Suite 2000
Indianapolis, IN. 46202
(317)278-0552
tracwilk@iu.edu

Manuscript Word Count: 2010

ABSTRACT

A Cross-Sectional Study of Predatory Publishing Emails Received by Career Development Grant Awardees

Objective: To investigate the scope of academic spam emails amongst career development grant awardees and factors associated with the amount of time spent addressing them.

Design: A cross-sectional survey of career development grant investigators via anonymous online survey was conducted. In addition to demographic and professional information, we asked investigators to report the number of academic spam emails received each day, how they determined whether these emails were spam, and time they spent per day addressing them. We used bivariate analysis to assess factors associated with the amount of time spent on academic spam emails.

Setting: An online survey sent via email on three separate occasions between Nov-December, 2016.

Participants: All National Institute of Health career development awardees funded in the 2015 fiscal year.

Main Outcome Measures: Factors associated with the amount of time spent addressing academic spam emails.

Results: A total of 3492 surveys were emailed, of which 206 (5.9%) were returned as undeliverable and 96 (2.7%) reported an out-of-office message; our overall response rate was 22.3% (n=733). All respondents reported receiving academic spam emails, with the majority (54.4%) receiving between 1 and 10 per day and spending between 1 and 10 minutes each day evaluating them. The amount of time respondents reported spending on academic spam emails was associated with the number of peer-reviewed journal articles authored ($p<0.001$), a history of publishing in open access format ($p<0.01$), the total number of academic spam emails received ($p<0.001$), and a feeling of having missed opportunities due to ignoring these emails ($p=0.04$).

Conclusions: Academic spam emails are a common distraction for career development grantees that may impact faculty productivity. There is an urgent need to mitigate this growing problem.

Keywords: Open Access Publishing, Predatory Journals, Publishing, Time Management

Strengths and Limitations of This Study

- This is the first study to describe the scope of academic spam emails amongst career development grantees
- The survey was distributed by email, thus could have been perceived as a spam email by recipients

- The survey only included recipients of National Institutes of Health funding, and did not include those that have applied for funding and were not successful

INTRODUCTION

In the last 10 years a subset of "predatory" publishers has been able to flourish within the $10.5 billion per year market. .[1-4]

These "predatory" publishers send unsolicited emails that request manuscript submissions, offer rapid review, and use publication fees (rather than traditional peer or editorial review) as criteria to accept articles.[5-7] As a result of this model, there have been documented instances of accepting flawed manuscripts or fake editorial board members.[8-11] Because academic research faculty career success is often centered on their publication and funding record, their ability to obtain funding to conduct research is intimately linked to their success in publishing their work. Junior academic faculty are particularly focused on opportunities to publish given the importance tied to their promotion. They may, therefore, be most susceptible to academic spam emails (ASEs) from predatory publishers, which contain unsolicited requests for publishing manuscripts, presentations at organization meetings, and memberships on editorial boards.

Previous studies have examined the quantity and quality of predatory publishers, as well as characteristics of authors who publish in them.[10, 12-19] However, to our knowledge, no study has described the prevalence of this phenomenon amongst researchers early in their career or quantified the time spent on ASEs. Therefore, we sought to examine factors associated with the amount of time spent addressing academic spam emails by career development grantees.

METHODS

Study Design and Data Source

We obtained the contact information for all National Institution of Health (NIH) K-awardees in the fiscal year of 2015 via a Freedom of Information Act request. Information of federal grant awardees is a standard report issued by the NIH Freedom of Information Office that is available online.[20] We then created and emailed an online survey using Survey Monkey[21] consisting of 14 questions to the correspondent/recipient on record of each K-award.

The survey was self-structured by the authors and piloted amongst 10 faculty at Indiana University School of Medicine that were not current K-awardees to assure clarity of questions and adequate response options. If a question was not clear or the answer choices not sufficient, the survey was edited before it was sent to the next faculty member. This process was repeated until there were no additional feedback suggestions. Validity testing on the survey was not done prior to the surveys being sent. The survey was sent a total of 3 times over a 2 month period (November-December) in 2016 to 3492 NIH K-awardees. The email was generated from a valid, personal email account of one of the study investigators (TAW).

The survey requested information regarding basic demographic (e.g., gender) and professional information (e.g., academic job title, year of terminal degree, and publication history). Survey respondents were also asked several questions about their experience with academic spam emails—which was described to survey respondents as an "unsolicited email requesting articles/editorial, conference presentations or editorial membership." Subsequent questions asked respondents to report the number of academic spam emails they received daily, how they determined whether these emails were spam,

the amount of time spent on these emails, whether their email account had a spam filter, and if they ever felt like they missed opportunities due to ignoring these emails.

Data Analysis

We summarized responses to survey questions using descriptive statistics and then assessed factors associated with the daily amount of time spent on these emails (re-categorized as none, 1-10 minutes and greater than 10 minutes) using chi-squared bivariate analysis. We removed respondents with missing data (n=68) or who replied "not sure" (n=4) to the question about time spent daily addressing academic spam emails from the bivariate analysis. Respondents removed from this analysis were less likely to have more than 20 publications, but otherwise did not differ from those included in the sample by key study variables. All analysis was done using SAS Version 9.4 (SAS Institute INC, Cary, NC).

Ethical Approval

The Indiana University IRB approved this study with a waiver of consent due to the lack of identifiable information being obtained in the anonymous survey responses.

Patient and Public Involvement

For this study, there was no patient involvement; however, we collected data through a survey of NIH grantees available from a public database.[20]

RESULTS

A total of 3492 surveys were emailed, of which 206 (5.9%) were immediately returned as undeliverable; 96 (2.7%) of these 3286 emails had an out-of-office message automatically in response to the inquiry. Of the emails sent and received successfully, 733 (22.3%) surveys were completed by the end of the study period and used for final

analysis. A slight majority were female (n=399; 54.4%). Over two-thirds (65.5%) of respondents were assistant professors working at an academic center; most (71.9%) received their NIH-K award funding in the 4 years preceding the survey (Table 1).

Table 1. Description of Survey Respondents*

Variable	Total N (%)
Total Reponses	733 (100)
Demographics and Professional Information	
Gender	
Female	399 (54.4)
Male	322 (43.9)
Other/Prefer Not to Answer	12 (1.6)
Academic Job Title (n=731)	
Assistant Professor	479 (65.5)
Associate Professor	120 (16.4)
Professor	49 (6.7)
Non Tenure-Track	57 (7.8)
Other	26 (3.6)
First Year to Receive NIH Funding (n=729)	
2011-2016	524 (71.9)
2010-2000	169 (23.3)
Prior to 2000	36 (4.9)
Number of Academic Publications (n=732)	
1 to 10	55 (7.5)
11 to 20	181 (24.7)
Greater than 20	496 (67.8)
Previously Published in Open Access Format (n=730)	
Yes	489 (67.0)
No	195 (26.7)
Not Sure	46 (6.3)
Year Terminal Degree Obtained (n=729)	
2010-2016	137 (18.8)
2005-2009	307 (42.1)
2000-2004	202 (27.7)
Prior to 2000	83 (11.4)
SPAM Filter on Email (n=729)	
Yes	651 (89.3)
No	36 (4.9)

Not Sure	42 (5.8)
Experience with Academic Spam Emails	
Estimated Number of Spam Emails/Day (n=724)	
1 to 10	398 (54.5)

Table 1. Description of Survey Respondents (continued)

Variable	Total N (%)
11 to 20	223 (30.6)
Greater than 20	103 (14.1)
Not Sure	6 (0.8)
Time Spent Per Day Reading/Sorting Spam Emails (n=665)	
None, delete them all without reading	132 (19.9)
1 to 10 minutes	419 (63.0)
11 to 20 minutes	89 (13.4)
More than 20 minutes	21 (3.2)
Not Sure	4 (0.6)
History of Responding to Academic Spam Email (n=729)	
Yes	226 (31.0)
No	503 (69.0)
Concern of Missed Opportunities (n=731)	
Yes	58 (7.9)
No	575 (78.7)
Not Sure	98 (13.4)
Methods used to Deem Emails Academic Spam[a]	
Consider them all Spam	274 (37.4)
Ask a Colleague	123 (16.8)
Typos in text/name	398 (54.5)
Don't Recognize Journal	601 (82.0)
Address listed is not in US	308 (42.0)
Requesting Fee to Publish	313 (42.7)
Consult the Internet	239 (32.6)

a. Respondents were instructed to select all that apply.

Note. Percentages may not sum to 100 due to rounding, total n differs if a question was left blank

With respect to number of ASEs received each day, every respondent reported receiving at least one academic spam email in the previous week (data not shown). Over half (54.5%) reported receiving between 1-10 academic spam emails per day, 30.6% reported receiving between 11 and 20 emails per day, and 14.1% reported receiving more than 20 emails per day (Table 1). When asked how much time was spent on these emails in a day, 80% of the respondents reported spending at least some amount of time during the day addressing these emails, with 63% of respondents reported spending between 1 and 10 minutes of their day reading, sorting, and determining what to do with potential ASEs. The methods of how they determined an email to be academic spam were diverse, with an unknown journal name being the most frequently described method (82% of respondents). More than a one third of respondents (31%) reported that they had responded to an ASE in the past.

On bivariate analysis, when examining variables associated with the amount of time spent on academic spam emails in a day (Table 2), we found that neither gender (p-value=0.37) nor academic job title (p-value=0.36) were associated with the time spent addressing academic spam emails. Interestingly, the time spent on emails was also not associated with reporting having a SPAM filter (p=0.99).

Table 2. Factors associated with time spent on emails

	Time Spent on Emails Per Day (%)			
	None	1-10min	>10min	p-value
Demographics and Professional Information				
Gender				0.37
Female	19.9	66.0	14.0	
Male	20.1	60.4	19.5	
Other/Prefer not to answer	16.7	58.3	25.0	
Academic Job Title				

Assistant Professor	19.5	65.4	15.1	0.36
Associate Professor	20.7	57.8	21.6	
Professor	28.6	51.0	20.4	
Non-Tenure Track	17.4	65.2	17.4	
Other	10.5	79.0	10.5	
Number of Academic Publications				
1 to 10	0.0	87.5	12.5	<0.001
11 to 20	0.0	80.8	19.2	
Greater than 20	26.9	56.8	16.3	
Previously Published in Open Access Format				
Yes	23.7	67.5	8.9	0.01
No	19.7	61.7	18.6	
Not Sure	9.3	65.1	25.6	
SPAM Filter on Email				
No	19.9	63.1	17.0	0.99
Yes	20.7	65.5	13.8	
Not Sure	20.0	65.0	15.0	
Experience with Academic Spam Emails				
Estimated Number of Spam Emails/Day				
1 to 10	20.3	72.9	6.8	<0.001
11 to 20	20.3	58.0	21.7	
Greater than 20	18.6	40.2	41.2	
Not Sure	0.0	100.0	0.0	
Concern for Missed Opportunities				
No	21.3	63.8	14.9	0.04
Yes	14.3	55.4	30.4	
Not Sure	15.9	65.9	18.2	

However, there was an inverse association (p-value <0.001) with the number of academic publications reported and the time spent addressing these emails, such that faculty with fewer publications reported spending more time per day reading potential ASEs and assessing their legitimacy. Among faculty with 10 or fewer publications, 87.5% spent between 1 and 10 minutes assessing these emails compared with those with 11-20 publications (80.8%) and greater than 20 (56.8%) publications, respectively. More than a quarter of respondents with greater than 20 publications spent no time assessing

these emails compared with 0% of respondents with fewer publications. There was also an association between the number of potential ASEs received in a day and the time spent on them (p-value <.001). Specifically, respondents who received more than 20 ASEs per day were more likely to report spending more than 10 minutes per day on these emails than respondents who received fewer emails.

Finally, we assessed whether respondents felt they might have missed opportunities in the past due because they mistakenly assumed legitimate emails were academic spam. We found that respondents who were concerned about missed opportunities were almost twice as likely to spend at least 10 minutes of their day assessing these emails than those who were not concerned (24.3% versus 12.8%; p-value=0.04).

DISCUSSION

This study sought to examine the scope of ASEs among NIH career development awardees and factors associated with the amount of daily time spent addressing them. We found that receiving ASEs was pervasive in this population. In fact, everyone who responded to the survey reported receiving academic spam, precluding us from examining demographic and job-related factors associated with the receipt of these emails. In addition, 80% of respondents reported using time during their day to address these emails.

We did examine these factors in relation to the amount of time respondents spent addressing the emails in any given day. Neither gender, academic rank, nor having a spam filter affected the time spent addressing these emails, but faculty with less than 20 publications and who felt like they might be missing opportunities did report spending

more time addressing these emails than their counterparts. More than 30% of respondents who felt that they might have missed opportunities for publication, presentation, or editorial service because they ignored an academic spam email reported spending more than 10 minutes every day reading and sorting academic spam. This suggests that academic faculty who are junior and/or who feel pressure to publish might be most susceptible to these types of predatory solicitations. In addition, the more ASEs received in a day was associated with more time spent addressing them.

While our study was focused on career development grantees that are generally in the early stages of their career, the burden of ASE's are not reserved for just these faculty. A 2017 study found that even senior faculty are affected by academic spam emails, making the case that efforts to control this phenomenon are needed throughout academia.[16] Therefore, whatever interventions or efforts are undertaken to combat this problem will need to include faculty at all stages and not just new or junior faculty.

As the number of journals and publishers continue to rise, we can expect the number of emails to do so as well. Given the focus of faculty to publish, it is likely that they will continue to spend time reading these emails to determine whether they are legitimate or not, when they could be working on projects, manuscripts, and grant proposals. Knowing that ASEs will likely not stop, this may indicate an additional piece of training faculty should receive and have resources given to determine whether or not to respond to unsolicited emails or any easy way to report them to help identify the senders as spam in the future.

Academic institutions could decide to value publications in "predatory" journals differently (in terms of promotion criteria) and thus deter faculty from pursuing those

opportunities, although that could be viewed as impacting academic freedom.[22] Further work is exploring how pervasive publication within these journals is ongoing and the larger question of how academia will perceive and evaluate these publications is yet to be answered.[23-25]

Our study has some limitations. First, we used the contact information for grant awardees provided by the NIH and did not determine individual contact information based on awardees name and institution separately. Thus, some email addresses were not working or valid at the time of our survey and may have represented an institutional contact and not the awardee themselves. Furthermore, our study sample was those with funded NIH grants, not including those that may have applied but were not funded. Second, there is a chance that our emailed survey was viewed by recipients as SPAM itself, and possibly ignored or deleted without a response. We used a valid personal email address with a survey link in the body of the email to solicit respondents, but that also prohibited us from knowing who had completed the survey already or who to follow-up with directly for not completing the survey. Third, the overall response rate was 22% and we did not require any question on the survey to be answered and thus had some missing data. However, respondents removed from the sample for missing data did not systematically differ from those included, except that they reported having fewer publications. This likely produced conservative results in our bivariate analysis. Nonetheless, our study quantifies the burden of time faculty spend addressing ASEs as career development grantees.

There are efforts underway to help researchers choose the right journal for their research which includes deciphering predatory journals from others.[18, 24, 26-31]

Furthermore, efforts at the university level to help improve email filters to block unsolicited emails can be refined and improved, given a majority of our respondents had SPAM filters in place that permitted these emails to be delivered. This may present challenges as to how to discern ASE's from other emails; however, certain patterns to emails or sender addresses could be used to flag these within a larger system. From a larger system perspective, the publishing industry can build upon their current communications to help differentiate from predatory journals for the receivers.

Acknowledgements: We would like to thank all the investigators that responded to this survey and Academic Librarian Jeffrey Beall who helped with the concept of this study. None of the authors have conflicts of interest or financial disclosures to report.

Competing Interests: None

Funding: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors

Author Contributions:
Dr. Wilkinson conceptualized and designed the study, distributed the survey, drafted the initial manuscript and approved the final manuscript as submitted.

Dr. Russell conceptualized and designed the study and approved the final manuscript as submitted

Dr. Bennett conceptualized and designed the study and approved the final manuscript as submitted

Dr. Cheng conceptualized and designed the study, carried out the analysis and approved the final manuscript as submitted.

Dr. Carroll contributed to the study design and data collection, reviewed and revised the manuscript, and approved of the final manuscript as submitted.

Data Sharing Statement: The survey, statistical code and dataset available by request to corresponding author, Dr. Tracey Wilkinson

References

1. Esposito P. The size of the open access market. The Scholarly Kitchen; 2014. http://scholarlykitchen.sspnet.org/2014/10/29/the-size-of-the-open-access-market.
2. Shen C, Björk B-C. 'Predatory' open access: a longitudinal study of article volumes and market characteristics. BMC Med. 2015;13(1):230.
3. Eriksson S, Helgesson G. The false academy: predatory publishing in science and bioethics. Med Health Care Philos. 2017;20(2):163-70.
4. Gasparyan AY, Yessirkepov M, Diyanova SN, Kitas GD. Publishing Ethics and Predatory Practices: A Dilemma for All Stakeholders of Science Communication. J Korean Med Sci. 2015;30(8):1010-6.
5. Cobey KD, Lalu MM, Skidmore B, Ahmadzai N, Grudniewicz A, Moher D. What is a predatory journal? A scoping review. F1000Research. 2018;7:1001-.
6. Beall J. Criteria for determining predatory open-access publishers. 2nd edition. Denver, CO: Scholarly Open Access; 2012. http://scholarlyoa.files.wordpress.com/2012/11/criteria-2012-2.pdf.
7. Beall J. Predatory publishers are corrupting open access. Nature. 2012;489(7415):179.
8. Daivs P. Open Access Publisher Accepts Nonsense Manuscript for Dollars 2009 [Available from: https://scholarlykitchen.sspnet.org/2009/06/10/nonsense-for-dollars/.
9. Butler D. Investigating journals: The dark side of publishing. Nature. 2013;495(7442):433-5.
10. Sorokowski P, Kulczycki E, Sorokowska A, Pisanski K. Predatory journals recruit fake editor. Nature. 2017;543(7646):481-3.
11. Bohannon J. Who's afraid of peer review. Science. 2013;342(6154).
12. Nwagwu WE, Ojemeni O. Penetration of Nigerian predatory biomedical open access journals 2007–2012: a bibliometric study. Learned Publishing. 2015;28(1):23-34.
13. Xia J. Predatory journals and their article publishing charges. Learned Publishing. 2015;28(1):69-74.
14. Xia J, Harmon JL, Connolly KG, Donnelly RM, Anderson MR, Howard HA. Who publishes in "predatory" journals? Journal of the Association for Information Science and Technology. 2015;66(7):1406-17.
15. Tin L, Ivana B, Biljana B, Ljubica IB, Dragan M, Dušan S. Predatory and fake scientific journals/publishers–a global outbreak with rising trend: a review. Geographica Pannonica. 2014;18(3):69-81.
16. Cobey K. Illegitimate journals scam even senior scientists. Nature. 2017;549(7670):7.
17. Bolshete P. Analysis of thirteen predatory publishers: a trap for eager-to-publish researchers. Curr Med Res Opin. 2018;34(1):157-62.
18. Memon AR. Predatory Journals Spamming for Publications: What Should Researchers Do? Sci Eng Ethics. 2018;24(5):1617-39.
19. 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(1):28.

20. Contact Information NIH-Supported PIs. National Institutes of Health-Freedom of Information Act Office 2015.
21. SurveyMonkey, Inc. San Mateo, California, USA.
22. Nelson C. Open Access and Academic Freedom: Inside Higher Ed; 2013 [Available from: https://www.insidehighered.com/views/2013/11/15/essay-impact-open-access-requirements-academic-freedom.
23. Pyne D. The Rewards of Predatory Publications at a Small Business School2017. 137-60 p.
24. Eriksson S, Helgesson G. Time to stop talking about 'predatory journals'. Learned Publishing. 2018;31(2):181-3.
25. Memon AR. Publish or perish: A sign of caution for authors to avoid predatory journals. J Pak Med Assoc. 2017;67(5):822.
26. Beall J. Best practices for scholarly authors in the age of predatory journals. Ann R Coll Surg Engl. 2016;98(2):77-9.
27. Laine C, Winker MA. Identifying predatory or pseudo-journals. Biochemia medica. 2017;27(2):285-91.
28. Memon AR. Research publications and education in Pakistani medical universities: Avoiding predatory journals and improving the quality of research. J Pak Med Assoc. 2017;67(6):830-3.
29. Teixeira da Silva JA, Tsigaris P. What Value Do Journal Whitelists and Blacklists Have in Academia? The Journal of Academic Librarianship. 2018;44(6):781-92.
30. Memon AR. How to respond to and what to do for papers published in predatory journals? Sci Ed. 2018;5(2):146-9.
31. Memon AR. ResearchGate and Impact Factor: A step further on predatory journals. J Pak Med Assoc. 2017;67(1):148-9.

STROBE Checklist

Manuscript Title "A Cross-Sectional Study of Predatory Publishing Emails Received by Career Development Grant Awardees: The ERASE Study

	Item No	Recommendation	Page/Line Number
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	Title Page
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	p. 2 Abstract
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	p.3-4
Objectives	3	State specific objectives, including any prespecified hypotheses	p.4, Line 5
Methods			
Study design	4	Present key elements of study design early in the paper	p.4, Methods
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	p. 4-5, Methods
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants	p. 4-5, Methods
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	p. 4-5, Methods
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	p. 4-5, Methods
Bias	9	Describe any efforts to address potential sources of bias	p. 12, Line 3
Study size	10	Explain how the study size was arrived at	p. 4-5, Methods
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	p. 4-5, Methods
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	p. 4-5, Methods
		(*b*) Describe any methods used to examine subgroups and interactions	N/A
		(*c*) Explain how missing data were addressed	p. 5, Data Analysis
		(*d*) If applicable, describe analytical methods taking account of sampling strategy	N/A
		(*e*) Describe any sensitivity analyses	N/A
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	p. 6, Results
		(b) Give reasons for non-participation at each stage	N/A
		(c) Consider use of a flow diagram	N/A
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential	p. 6 Results, Table 1

		confounders	
		(b) Indicate number of participants with missing data for each variable of interest	Table 1
Outcome data	15*	Report numbers of outcome events or summary measures	Table 1
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	N/A
		(*b*) Report category boundaries when continuous variables were categorized	Table 1
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	N/A
Discussion			
Key results	18	Summarise key results with reference to study objectives	p. 11, Discussion
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	p. 12, Line 3
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	p. 12-13
Generalisability	21	Discuss the generalisability (external validity) of the study results	p. 12-13
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	p. 13

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

A Cross-Sectional Study of Predatory Publishing Emails Received by Career Development Grant Awardees

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-027928.R3
Article Type:	Research
Date Submitted by the Author:	18-Apr-2019
Complete List of Authors:	Wilkinson, Tracey; Indiana University Department of Pediatrics, Children's Health Services Research Russell, Christopher; Children's Hospital of Los Angeles; Keck School of Medicine of the University of Southern California Bennett, William; Indiana University School of Medicine, Pediatrics, Division of Gastroenterology Cheng, Erika; Indiana University Department of Pediatrics, Children's Health Services Research Carroll, Aaron; Indiana University Department of Pediatrics
Primary Subject Heading:	Medical publishing and peer review
Secondary Subject Heading:	Medical education and training
Keywords:	publishing, time management, predatory journals, open access publishing

A Cross-Sectional Study of Predatory Publishing Emails Received by Career Development Grant Awardees

Tracey A. Wilkinson, MD, MPH[1*], Christopher J. Russell, MD, MS,[2] William E. Bennett, Jr, MD, MS[1], Erika R. Cheng, PhD, MPA[1], Aaron E. Carroll, MD, MS[3]

[1]Assistant Professor, Department of Pediatrics, Indiana University School of Medicine, Indianapolis, IN
[2]Assistant Professor, Department of Pediatrics, Keck School of Medicine, University of Southern California, Los Angeles, CA, Division of Hospital Medicine, Children's Hospital Los Angeles, Los Angeles, CA, United States
[3]Professor, Department of Pediatrics, Indiana University School of Medicine, Indianapolis, IN

* Corresponding Author:
Tracey A. Wilkinson
410 W. 10th Street, Suite 2000
Indianapolis, IN. 46202
(317)278-0552
tracwilk@iu.edu

Manuscript Word Count: 2010

ABSTRACT

A Cross-Sectional Study of Predatory Publishing Emails Received by Career Development Grant Awardees

Objective: To investigate the scope of academic spam emails amongst career development grant awardees and factors associated with the amount of time spent addressing them.

Design: A cross-sectional survey of career development grant investigators via anonymous online survey was conducted. In addition to demographic and professional information, we asked investigators to report the number of academic spam emails received each day, how they determined whether these emails were spam, and time they spent per day addressing them. We used bivariate analysis to assess factors associated with the amount of time spent on academic spam emails.

Setting: An online survey sent via email on three separate occasions between Nov-December, 2016.

Participants: All National Institute of Health career development awardees funded in the 2015 fiscal year.

Main Outcome Measures: Factors associated with the amount of time spent addressing academic spam emails.

Results: A total of 3492 surveys were emailed, of which 206 (5.9%) were returned as undeliverable and 96 (2.7%) reported an out-of-office message; our overall response rate was 22.3% (n=733). All respondents reported receiving academic spam emails, with the majority (54.4%) receiving between 1 and 10 per day and spending between 1 and 10 minutes each day evaluating them. The amount of time respondents reported spending on academic spam emails was associated with the number of peer-reviewed journal articles authored ($p<0.001$), a history of publishing in open access format ($p<0.01$), the total number of academic spam emails received ($p<0.001$), and a feeling of having missed opportunities due to ignoring these emails ($p=0.04$).

Conclusions: Academic spam emails are a common distraction for career development grantees that may impact faculty productivity. There is an urgent need to mitigate this growing problem.

Keywords: Open Access Publishing, Predatory Journals, Publishing, Time Management

Strengths and Limitations of This Study

- This is the first study to describe the scope of academic spam emails amongst career development grantees
- The survey was distributed by email, thus could have been perceived as a spam email by recipients

- The survey only included recipients of National Institutes of Health funding, and did not include those that have applied for funding and were not successful

INTRODUCTION

In the last 10 years, a subset of "predatory" publishers has been able to flourish within the $10.5 billion per year market. [1-4] These "predatory" publishers send unsolicited emails that request manuscript submissions, offer rapid review, and use publication fees (rather than traditional peer or editorial review) as criteria to accept articles.[5-7] As a result of this model, there have been documented instances of accepting flawed manuscripts or fake editorial board members.[8-11] Because academic research faculty career success is often centered on their publication and funding record, their ability to obtain funding to conduct research is intimately linked to their success in publishing their work. Junior academic faculty are particularly focused on opportunities to publish given the importance tied to their promotion. They may, therefore, be most susceptible to academic spam emails (ASEs) from predatory publishers, which contain unsolicited requests for publishing manuscripts, presentations at organization meetings, and memberships on editorial boards.

Previous studies have examined the quantity and quality of predatory publishers, as well as characteristics of authors who publish in them.[10, 12-19] However, to our knowledge, no study has described the prevalence of this phenomenon amongst researchers early in their career or quantified the time spent on ASEs. Therefore, we sought to examine factors associated with the amount of time spent addressing academic spam emails by career development grantees.

METHODS

Study Design and Data Source

We obtained the contact information for all National Institution of Health (NIH) K-awardees in the fiscal year of 2015 via a Freedom of Information Act request. Information of federal grant awardees is a standard report issued by the NIH Freedom of Information Office that is available online.[20] We then created and emailed an online survey using Survey Monkey[21] consisting of 14 questions to the correspondent/recipient on record of each K-award.

The survey was conducted using a self-structured questionnaire, which was piloted amongst 10 faculty at Indiana University School of Medicine that were not current K-awardees to assure clarity of questions and adequate response options. If a question was not clear or the answer choices not sufficient, the survey was edited before it was sent to the next faculty member. This process was repeated until there were no additional feedback suggestions. Validity testing on the survey (prior to the surveys being sent to the study participants) was not done. The survey was sent a total of 3 times over a 2 month period (November-December) in 2016 to 3492 NIH K-awardees. The email was generated from a valid, personal email account of one of the study investigators (TAW).

The survey requested information regarding basic demographic (e.g., gender) and professional information (e.g., academic job title, year of terminal degree, and publication history). Survey respondents were also asked several questions about their experience with academic spam emails—which was described to survey respondents as an "unsolicited email requesting articles/editorial, conference presentations or editorial membership." Subsequent questions asked respondents to report the number of academic spam emails they received daily, how they determined whether these emails were spam,

the amount of time spent on these emails, whether their email account had a spam filter, and if they ever felt like they missed opportunities due to ignoring these emails.

Data Analysis

We summarized responses to survey questions using descriptive statistics and then assessed factors associated with the daily amount of time spent on these emails (re-categorized as none, 1-10 minutes and greater than 10 minutes) using chi-squared bivariate analysis. We removed respondents with missing data (n=68) or who replied "not sure" (n=4) to the question about time spent daily addressing academic spam emails from the bivariate analysis. Respondents removed from this analysis were less likely to have more than 20 publications, but otherwise did not differ from those included in the sample by key study variables. All analysis was done using SAS Version 9.4 (SAS Institute INC, Cary, NC).

Ethical Approval

The Indiana University IRB approved this study with a waiver of consent due to the lack of identifiable information being obtained in the anonymous survey responses.

Patient and Public Involvement

For this study, there was no patient involvement; however, we collected data through a survey of NIH grantees available from a public database.[20]

RESULTS

A total of 3492 surveys were emailed, of which 206 (5.9%) were immediately returned as undeliverable; 96 (2.7%) of these 3286 emails had an out-of-office message automatically in response to the inquiry. Of the emails sent and received successfully, 733 (22.3%) surveys were completed by the end of the study period and used for final

analysis. A slight majority were female (n=399; 54.4%). Over two-thirds (65.5%) of respondents were assistant professors working at an academic center; most (71.9%) received their NIH-K award funding in the 4 years preceding the survey (Table 1).

Table 1. Description of Survey Respondents*

Variable	Total N (%)
Total Reponses	733 (100)
Demographics and Professional Information	
Gender	
Female	399 (54.4)
Male	322 (43.9)
Other/Prefer Not to Answer	12 (1.6)
Academic Job Title (n=731)	
Assistant Professor	479 (65.5)
Associate Professor	120 (16.4)
Professor	49 (6.7)
Non Tenure-Track	57 (7.8)
Other	26 (3.6)
First Year to Receive NIH Funding (n=729)	
2011-2016	524 (71.9)
2010-2000	169 (23.3)
Prior to 2000	36 (4.9)
Number of Academic Publications (n=732)	
1 to 10	55 (7.5)
11 to 20	181 (24.7)
Greater than 20	496 (67.8)
Previously Published in Open Access Format (n=730)	
Yes	489 (67.0)
No	195 (26.7)
Not Sure	46 (6.3)
Year Terminal Degree Obtained (n=729)	
2010-2016	137 (18.8)
2005-2009	307 (42.1)
2000-2004	202 (27.7)
Prior to 2000	83 (11.4)
SPAM Filter on Email (n=729)	
Yes	651 (89.3)
No	36 (4.9)

Not Sure	42 (5.8)
Experience with Academic Spam Emails	
Estimated Number of Spam Emails/Day (n=724)	
1 to 10	398 (54.5)

Table 1. Description of Survey Respondents (continued)

Variable	Total N (%)
11 to 20	223 (30.6)
Greater than 20	103 (14.1)
Not Sure	6 (0.8)
Time Spent Per Day Reading/Sorting Spam Emails (n=665)	
None, delete them all without reading	132 (19.9)
1 to 10 minutes	419 (63.0)
11 to 20 minutes	89 (13.4)
More than 20 minutes	21 (3.2)
Not Sure	4 (0.6)
History of Responding to Academic Spam Email (n=729)	
Yes	226 (31.0)
No	503 (69.0)
Concern of Missed Opportunities (n=731)	
Yes	58 (7.9)
No	575 (78.7)
Not Sure	98 (13.4)
Methods used to Deem Emails Academic Spam[a]	
Consider them all Spam	274 (37.4)
Ask a Colleague	123 (16.8)
Typos in text/name	398 (54.5)
Don't Recognize Journal	601 (82.0)
Address listed is not in US	308 (42.0)
Requesting Fee to Publish	313 (42.7)
Consult the Internet	239 (32.6)

a. Respondents were instructed to select all that apply.

Note. Percentages may not sum to 100 due to rounding, total n differs if a question was left blank

With respect to number of ASEs received each day, every respondent reported receiving at least one academic spam email in the previous week (data not shown). Over half (54.5%) reported receiving between 1-10 academic spam emails per day, 30.6% reported receiving between 11 and 20 emails per day, and 14.1% reported receiving more than 20 emails per day (Table 1). When asked how much time was spent on these emails in a day, 80% of the respondents reported spending at least some amount of time during the day addressing these emails, with 63% of respondents reported spending between 1 and 10 minutes of their day reading, sorting, and determining what to do with potential ASEs. The methods of how they determined an email to be academic spam were diverse, with an unknown journal name being the most frequently described method (82% of respondents). More than a one third of respondents (31%) reported that they had responded to an ASE in the past.

On bivariate analysis, when examining variables associated with the amount of time spent on academic spam emails in a day (Table 2), we found that neither gender (p-value=0.37) nor academic job title (p-value=0.36) were associated with the time spent addressing academic spam emails. Interestingly, the time spent on emails was also not associated with reporting having a SPAM filter (p=0.99).

Table 2. Factors associated with time spent on emails

	Time Spent on Emails Per Day (%)			
	None	1-10min	>10min	p-value
Demographics and Professional Information				
Gender				0.37
Female	19.9	66.0	14.0	
Male	20.1	60.4	19.5	
Other/Prefer not to answer	16.7	58.3	25.0	
Academic Job Title				

Assistant Professor	19.5	65.4	15.1	0.36
Associate Professor	20.7	57.8	21.6	
Professor	28.6	51.0	20.4	
Non-Tenure Track	17.4	65.2	17.4	
Other	10.5	79.0	10.5	
Number of Academic Publications				
1 to 10	0.0	87.5	12.5	<0.001
11 to 20	0.0	80.8	19.2	
Greater than 20	26.9	56.8	16.3	
Previously Published in Open Access Format				
Yes	23.7	67.5	8.9	0.01
No	19.7	61.7	18.6	
Not Sure	9.3	65.1	25.6	
SPAM Filter on Email				
No	19.9	63.1	17.0	0.99
Yes	20.7	65.5	13.8	
Not Sure	20.0	65.0	15.0	
Experience with Academic Spam Emails				
Estimated Number of Spam Emails/Day				
1 to 10	20.3	72.9	6.8	<0.001
11 to 20	20.3	58.0	21.7	
Greater than 20	18.6	40.2	41.2	
Not Sure	0.0	100.0	0.0	
Concern for Missed Opportunities				
No	21.3	63.8	14.9	0.04
Yes	14.3	55.4	30.4	
Not Sure	15.9	65.9	18.2	

However, there was an inverse association (p-value <0.001) with the number of academic publications reported and the time spent addressing these emails, such that faculty with fewer publications reported spending more time per day reading potential ASEs and assessing their legitimacy. Among faculty with 10 or fewer publications, 87.5% spent between 1 and 10 minutes assessing these emails compared with those with 11-20 publications (80.8%) and greater than 20 (56.8%) publications, respectively. More than a quarter of respondents with greater than 20 publications spent no time assessing

these emails compared with 0% of respondents with fewer publications. There was also an association between the number of potential ASEs received in a day and the time spent on them (p-value <.001). Specifically, respondents who received more than 20 ASEs per day were more likely to report spending more than 10 minutes per day on these emails than respondents who received fewer emails.

Finally, we assessed whether respondents felt they might have missed opportunities in the past due because they mistakenly assumed legitimate emails were academic spam. We found that respondents who were concerned about missed opportunities were almost twice as likely to spend at least 10 minutes of their day assessing these emails than those who were not concerned (24.3% versus 12.8%; p-value=0.04).

DISCUSSION

This study sought to examine the scope of ASEs among NIH career development awardees and factors associated with the amount of daily time spent addressing them. We found that receiving ASEs was pervasive in this population. In fact, everyone who responded to the survey reported receiving academic spam, precluding us from examining demographic and job-related factors associated with the receipt of these emails. In addition, 80% of respondents reported using time during their day to address these emails.

We examined demographic and job-related factors in relation to the amount of time respondents spent addressing the emails in any given day. Neither gender, academic rank, nor having a spam filter affected the time spent addressing these emails, but faculty with less than 20 publications and who felt like they might be missing opportunities did

report spending more time addressing these emails than their counterparts. More than 30% of respondents who felt that they might have missed opportunities for publication, presentation, or editorial service because they ignored an academic spam email reported spending more than 10 minutes every day reading and sorting academic spam. This suggests that academic faculty who are junior and/or who feel pressure to publish might be most susceptible to these types of predatory solicitations. In addition, the more ASEs received in a day was associated with more time spent addressing them.

While our study was focused on career development grantees that are generally in the early stages of their career, the burden of ASE's is not limited to these faculty. In 2017, a study found that even senior faculty are affected by academic spam emails, making the case that efforts to control this phenomenon are needed throughout academia.[16] Therefore, whatever interventions or efforts are undertaken to combat this problem will need to include faculty at all stages and not just new or junior faculty.

As the number of journals and publishers continue to rise, we can expect the number of emails to do so as well. Given the focus of faculty to publish, it is likely that they will continue to spend time reading these emails to determine whether they are legitimate or not, when they could be working on projects, manuscripts, and grant proposals. Knowing that ASEs will likely not stop, this may indicate an additional piece of training faculty should receive and have resources given to determine whether or not to respond to unsolicited emails or any easy way to report them to help identify the senders as spam in the future.

Academic institutions could decide to value publications in "predatory" journals differently (in terms of promotion criteria) and thus deter faculty from pursuing those

opportunities, although that could be viewed as impacting academic freedom.[22] Additional research exploring how pervasive publication within these journals is occurring amongst academicians is underway; however, larger question of how academia will perceive and evaluate these publications within the promotion and tenure setting is yet to be answered.[23-25]

Our study has some limitations. First, we used the contact information for grant awardees provided by the NIH and did not determine individual contact information based on awardees name and institution separately. Thus, some email addresses were not working or valid at the time of our survey and may have represented an institutional contact and not the awardee themselves. Furthermore, our study sample was those with funded NIH grants, not including those that may have applied but were not funded. Second, there is a chance that our emailed survey was viewed by recipients as SPAM itself, and possibly ignored or deleted without a response. We used a valid personal email address with a survey link in the body of the email to solicit respondents, but that also prohibited us from knowing who had completed the survey already or who to follow-up with directly for not completing the survey. Third, the overall response rate was 22% and we did not require any question on the survey to be answered and thus had some missing data. However, respondents removed from the sample for missing data did not systematically differ from those included, except that they reported having fewer publications. This likely produced conservative results in our bivariate analysis.

Nonetheless, our study quantifies the burden of time faculty spend addressing ASEs as career development grantees and can provide groundwork for further studies examining the burden of ASE's amongst faculty.There are efforts underway to help

researchers choose the right journal for their research which includes deciphering predatory journals from others.[18, 24, 26-31] Furthermore, efforts at the university level to help improve email filters to block unsolicited emails can be refined and improved, given a majority of our respondents had SPAM filters in place that permitted these emails to be delivered. This may present challenges as to how to discern ASE's from other emails; however, certain patterns to emails or sender addresses could be used to flag these within a larger system. From a larger system perspective, the publishing industry can build upon their current communications to help differentiate from predatory journals for the receivers.

Acknowledgements: We would like to thank all the investigators that responded to this survey and Academic Librarian Jeffrey Beall who helped with the concept of this study. None of the authors have conflicts of interest or financial disclosures to report.

Competing Interests: None

Funding: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors

Author Contributions:
Dr. Wilkinson conceptualized and designed the study, distributed the survey, drafted the initial manuscript and approved the final manuscript as submitted.

Dr. Russell conceptualized and designed the study and approved the final manuscript as submitted

Dr. Bennett conceptualized and designed the study and approved the final manuscript as submitted

Dr. Cheng conceptualized and designed the study, carried out the analysis and approved the final manuscript as submitted.

Dr. Carroll contributed to the study design and data collection, reviewed and revised the manuscript, and approved of the final manuscript as submitted.

Data Sharing Statement: The survey, statistical code and dataset available by request to corresponding author, Dr. Tracey Wilkinson

References

1. Esposito P. The size of the open access market. The Scholarly Kitchen; 2014. http://scholarlykitchen.sspnet.org/2014/10/29/the-size-of-the-open-access-market.
2. Shen C, Björk B-C. 'Predatory' open access: a longitudinal study of article volumes and market characteristics. BMC Med. 2015;13(1):230.
3. Eriksson S, Helgesson G. The false academy: predatory publishing in science and bioethics. Med Health Care Philos. 2017;20(2):163-70.
4. Gasparyan AY, Yessirkepov M, Diyanova SN, Kitas GD. Publishing Ethics and Predatory Practices: A Dilemma for All Stakeholders of Science Communication. J Korean Med Sci. 2015;30(8):1010-6.
5. Cobey KD, Lalu MM, Skidmore B, Ahmadzai N, Grudniewicz A, Moher D. What is a predatory journal? A scoping review. F1000Research. 2018;7:1001-.
6. Beall J. Criteria for determining predatory open-access publishers. 2nd edition. Denver, CO: Scholarly Open Access; 2012. http://scholarlyoa.files.wordpress.com/2012/11/criteria-2012-2.pdf.
7. Beall J. Predatory publishers are corrupting open access. Nature. 2012;489(7415):179.
8. Daivs P. Open Access Publisher Accepts Nonsense Manuscript for Dollars 2009 [Available from: https://scholarlykitchen.sspnet.org/2009/06/10/nonsense-for-dollars/.
9. Butler D. Investigating journals: The dark side of publishing. Nature. 2013;495(7442):433-5.
10. Sorokowski P, Kulczycki E, Sorokowska A, Pisanski K. Predatory journals recruit fake editor. Nature. 2017;543(7646):481-3.
11. Bohannon J. Who's afraid of peer review. Science. 2013;342(6154).
12. Nwagwu WE, Ojemeni O. Penetration of Nigerian predatory biomedical open access journals 2007–2012: a bibliometric study. Learn Publ. 2015;28(1):23-34.
13. Xia J. Predatory journals and their article publishing charges. Learn Publ. 2015;28(1):69-74.
14. Xia J, Harmon JL, Connolly KG, Donnelly RM, Anderson MR, Howard HA. Who publishes in "predatory" journals? J Assoc Inf Sci Technol. 2015;66(7):1406-17.
15. Tin L, Ivana B, Biljana B, Ljubica IB, Dragan M, Dušan S. Predatory and fake scientific journals/publishers–a global outbreak with rising trend: a review. Geographica Pannonica. 2014;18(3):69-81.
16. Cobey K. Illegitimate journals scam even senior scientists. Nature. 2017;549(7670):7.
17. Bolshete P. Analysis of thirteen predatory publishers: a trap for eager-to-publish researchers. Curr Med Res Opin. 2018;34(1):157-62.
18. Memon AR. Predatory Journals Spamming for Publications: What Should Researchers Do? Science and engineering ethics. 2018;24(5):1617-39.
19. 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(1):28.
20. Contact Information NIH-Supported PIs. National Institutes of Health-Freedom of Information Act Office 2015.

21. SurveyMonkey, Inc. San Mateo, California, USA.
22. Nelson C. Open Access and Academic Freedom: Inside Higher Ed; 2013 [Available from: https://www.insidehighered.com/views/2013/11/15/essay-impact-open-access-requirements-academic-freedom.
23. Pyne D. The Rewards of Predatory Publications at a Small Business School2017. 137-60 p.
24. Eriksson S, Helgesson G. Time to stop talking about 'predatory journals'. Learned Publishing. 2018;31(2):181-3.
25. Memon AR. Publish or perish: A sign of caution for authors to avoid predatory journals. J Pak Med Assoc. 2017;67(5):822.
26. Beall J. Best practices for scholarly authors in the age of predatory journals. Ann R Coll Surg Engl. 2016;98(2):77-9.
27. Laine C, Winker MA. Identifying predatory or pseudo-journals. Biochemia medica. 2017;27(2):285-91.
28. Memon AR. Research publications and education in Pakistani medical universities: Avoiding predatory journals and improving the quality of research. J Pak Med Assoc. 2017;67(6):830-3.
29. Teixeira da Silva JA, Tsigaris P. What Value Do Journal Whitelists and Blacklists Have in Academia? J Academ Lib. 2018;44(6):781-92.
30. Memon AR. How to respond to and what to do for papers published in predatory journals? Sci Ed. 2018;5(2):146-9.
31. Memon AR. ResearchGate and Impact Factor: A step further on predatory journals. J Pak Med Assoc. 2017;67(1):148-9.

STROBE Checklist

Manuscript Title "A Cross-Sectional Study of Predatory Publishing Emails Received by Career Development Grant Awardees: The ERASE Study

	Item No	Recommendation	Page/Line Number
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	Title Page
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	p. 2 Abstract
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	p.3-4
Objectives	3	State specific objectives, including any prespecified hypotheses	p.4, Line 5
Methods			
Study design	4	Present key elements of study design early in the paper	p.4, Methods
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	p. 4-5, Methods
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants	p. 4-5, Methods
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	p. 4-5, Methods
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	p. 4-5, Methods
Bias	9	Describe any efforts to address potential sources of bias	p. 12, Line 3
Study size	10	Explain how the study size was arrived at	p. 4-5, Methods
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	p. 4-5, Methods
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	p. 4-5, Methods
		(*b*) Describe any methods used to examine subgroups and interactions	N/A
		(*c*) Explain how missing data were addressed	p. 5, Data Analysis
		(*d*) If applicable, describe analytical methods taking account of sampling strategy	N/A
		(*e*) Describe any sensitivity analyses	N/A
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	p. 6, Results
		(b) Give reasons for non-participation at each stage	N/A
		(c) Consider use of a flow diagram	N/A
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential	p. 6 Results, Table 1

		confounders	
		(b) Indicate number of participants with missing data for each variable of interest	Table 1
Outcome data	15*	Report numbers of outcome events or summary measures	Table 1
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	N/A
		(*b*) Report category boundaries when continuous variables were categorized	Table 1
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	N/A
Discussion			
Key results	18	Summarise key results with reference to study objectives	p. 11, Discussion
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	p. 12, Line 3
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	p. 12-13
Generalisability	21	Discuss the generalisability (external validity) of the study results	p. 12-13
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	p. 13

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.