PubMed-indexed duplicate publications in the last decade, 1996-2006

To the Editor: I have read with great interest the recently published editorial on the policy of your prestigious journal concerning duplicate publication.¹ It motivated me to perform an analysis of the PubMed-indexed publications from the last decade that were assigned the NLM Publication Type "Duplicate publication [PT]", aiming to clarify to the readers and authors some of what they should know about duplicate publication. A search strategy was undertaken including "Duplicate publication [PT]" directed to PubMed within a single hour limit (on April, 25th, 2007), limited to citations "published in the last 10 years" from the aforementioned date. The number of all PubMed-listed biomedical studies published in the last decade totaled 5 553 528 citations, of them 30 744 belonged to authors affiliated to Arab institutions (0.55%). Of the total 5 553 528 citations, 382 were of the "duplicate publication" type. The Medline format of the text file was then converted, using a locally designed visual basic program into a Microsoft Excel file, which eventually was captured as a new database query via SPSS software ending with 382 records in SPSS data file format. Because the search of duplicate publications in PubMed returns almost the double its actual number by showing both the original and the duplicate, the author read thoroughly all the abstracts of the 382 articles and looked meticulously into its contents. Moreover, the author sorted the fields of title and first author name in the data file to detect resemblance in the titles and authors' names in all studies. However, some duplicate publications have different but near titles

and a few have the same first author but a different number of co-authors. Finally, the actual number of duplicate publication was only 226 articles (0.004%) without any duplication in the SPSS data file. The reason behind exceeding the half could be explained by many factors. Some of articles appeared only as a single duplicate publication in PubMed because it was earlier and originally published in a non-indexed journal and then in the indexed one.² Also, there could be some delay or reluctance in notifying PubMed. Some of the 2006 duplicate publications might only show in 2007, which was not included in the search. Table 1 shows that the number of duplicate publications increased during the years 2002-2007 than in the first five years of the last decade. There was no significant change in the percentage of research on human subjects, articles in English, nor the place of journal of publication between the first and second five years of the last decade. There was a nonsignificant trend of increase in author number and also a nonsignificant decrease in the percentage of review, clinical trials and editorial articles of duplicate publication between the two groups of years. Only two articles from the total of 226 were retracted and two studies—one of them was retracted—were authored by a researcher affiliated to Arab institutions.

Duplicate publication is considered a type of self-plagiarism along with data fragmentation (i.e., salami-slicing), copyright infringement, and the practice of text recycling.³ Plagiarism is taking over ideas or words of another without acknowledging him or her. Using one's own written product in a new article without letting the reader to know the original source of the first product is self plagiarism.³ Duplicate publications occur be-

cause of the pressure to publish in a culture of "Publish or Perish", not understanding the seriousness and how duplicate publication could distort data aggregation in systemic reviews, the reluctance of academic leaders or journals to condemn the practice, or punish cases, not doing enough by journals to publish clear definitions and/or identify cases during peer review process.⁴ The length of peer review time and the uncertainty of accepting research for publication felt by authors could also be added to the aforementioned causes. Although the percentage of duplicate publications from the total (0.004%) in the last decade was very small, it does not give the true picture of the problem, especially for the Arab world where authors mostly publish in non-indexed journals. Mojan-Azzi et al⁵ developed an electronic search engine for redundancies to estimate the amount of duplicate publication in scientific journals. They applied their method to 22 433 articles published between 1997 and 2000 in 70 Medline-indexed ophthalmologic journals. Albeit they estimated that 1.39% of the publications were redundant, they regarded their estimate as the "tip of the iceberg".⁵

Because "no uniform standards and procedures for investigating duplicate publications are available", journal editors, suspecting duplicate publication, should work together, and invite independent reviewers to make the decision of duplicate publication.⁶ Journals could also "require that authors sign statement denying or disclosing overlapping" in their manuscripts, and "arrange with other journals to impose sanctions on authors of redundant articles".⁴ Moreover, there is a need to register planned and ongoing clinical trials and to change emphasis from quantity to quality by requiring applicants of post and grants.⁵

Table 1. Comparison of PubMed indexed articles of duplicate publication type in the last decade 1997-2006 (n=226).

Characteristics of duplicate publications		1997-2001(n=83)	2002-2006(n=143)	Total	Tests of significance, <i>P</i>
Number of authors (in 9 articles no author was mentioned)					
Single author	N	17	19	36	
	%	21.79	13.67	16.59	
4-5 authors	N	40	82	122	
	%	51.28	58.99	56.22	
5 or more	N	21	38	59	
	%	26.92	27.34	27.19	$\chi^2=2.46, 0.29$
Mean number of authors	Mean	4.26	4.45	4.15	F=0.4, 0.527
Clinical study (Human subjects)	N	70	106	176	
	%	84.34	74.13	77.88	$\chi^2=3.3, 0.096$
Article in English	N	69	122	191	
	%	83.13	85.31	84.51	$\chi^2=0.19, 0.71$
Type of publication					
Journal article	N	37	77	114	
	%	44.58	53.85	50.44	
Review	N	12	13	25	
	%	14.46	9.09	11.06	
Clinical trial	N	7	10	17	
	%	8.43	6.99	7.52	
Editorial	N	3	3	6	
	%	3.61	2.10	2.65	
Others	N	24	40	64	
	%	28.92	27.97	28.32	$\chi^2=2.83, 0.59$
Place of journal publication					
USA	N	35	63	98	
	%	42.17	44.06	43.36	
Elsewhere	N	48	80	128	
	%	57.83	55.94	56.64	$\chi^2=0.08, 0.89$
Authors affiliated to Arab institutions	N	1	1	2	
	%	1.2	0.7	0.9	Fisher exact not significant, <i>P</i> =1
Retracted publication					
	N	1	1	2	
	%	1.2	0.7	0.9	Fisher exact not significant, <i>P</i> =1

Responding to duplicate publications, journals should notify authors and impose restrictions on their future submission, inform authors' institutions and funding sources of the research, publish notices for the readers, and inform other journals and indexing services.⁴ Only in severe cases, is duplicate publication retracted,⁶ which is consistent with what the current study data shows of retraction of duplicate publications in only 0.5% of the cases.

Von Elm et al⁷ analyzed articles in systematic reviews in anesthesia and analgesia, 1989-2002, to investigate patterns of duplication and propose a decision tree for classification. They identified different patterns of duplications and found that of the duplicates, 64% had partly or completely different authors from the main article. Hence they concluded that authorship was an unreliable criteria of duplication.⁷ Therefore, the decision of banning the co-authors from submitting manuscripts to the *Annals of Saudi Medicine* would not be questionable only if it came after ensuring the similarity of the list of authors on the byline for both the main and duplicate articles. It could be difficult, especially when some of the regional journals still do not require all authors on the byline to sign an assignment of copyright. Moreover, some regional journals do not require assignment of copyright at all. The shared responsibility of all authors for their article is undeniable. However, there could be some odds of the innocence of a part of the co-authors. Then, the first or the corresponding author should take the responsibility of his misconduct. In the Poehlman case of data falsification,⁸ Poehlman⁹ took the sole responsibility for the falsification, and publicly exonerated his coauthors in his letter of apology.⁹

Dr. Mustafa Afifi

Department of Non-Communicable Diseases, Ministry of Health
P.O. Box 393, P.C 113
Muscat 113
Oman
T: +968 9903 5672
F: +968 2469 5480
afifidr@gmail.com

REFERENCES

1. Cathey JT, Aljurf M. Duplicate and fragmented publication: the policy of the Annals of Saudi Medicine. *Ann Saudi Med.* 2007;27(1):iii.
2. Hamza AH, Nasr MM, Deghady MM. The use of mitomycin-C for respiratory papillomas: clinical, histologic and biochemical correlation. *Saudi Med J.* 2006 ;27(4):572
3. Roig M. Avoiding plagiarism, self-plagiarism, and other questionable writing practices: A guide to ethical writing. Revised on-line version published in August, 2006. URL: <http://facpub.stjohns.edu/~roigm/plagiarism/Index.html> [accessed on 26 April 2006]
4. Yank V, Barnes D. Consensus and contention regarding redundant publications in clinical research: cross-sectional survey of editors and authors. *J Med Ethics.* 2003;29(2):109-14.
5. Mojon-Azzi SM, Jiang X, Wagner U, Mojon DS. Redundant publications in scientific ophthalmologic journals: the tip of the iceberg? *Ophthalmology.* 2004;111(5):863-6.
6. Hegyvary ST. What Every Author Should Know About Redundant and Duplicate Publication. *J Nurs Scholarsh.* 2005;37(40):295-297(3) URL: <http://www.blackwell-synergy.com/doi/pdf/10.1111/j.1547-5069.2005.00051.x?cookieSet=1> [accessed on 26 April 2006]
7. von Elm E, Poglía G, Walder B, Tramer MR. Different patterns of duplicate publication: an analysis of articles used in systematic reviews. *JAMA.* 2004;291(8):974-80.
8. Sox HC, Rennie D. Research misconduct, retraction, and cleansing the medical literature: lessons from the Poehlman case. *Ann Intern Med.* 2006;144(8):609-13.
9. Poehlman ET. Letter of apology for falsification of data. *Am J Physiol Endocrinol Metab.* 2005;289(2):E357.

sion. If they later convince us that their names were placed on the new submission without their knowledge, we may remove them from the banned list. As you say, they could be innocent. In searching for possible previous publication, we rely initially more on the title and abstract than on author names. When different authors appear on the new submission compared with the original publication this suggests that new authors were not involved in the study reported in the first published version of the manuscript, raising the issue of contributorship in addition to duplicate submission. Presumably their contribution to the study was not sufficient to warrant inclusion in the original list of authors. We have to assume that all authors on the new, duplicate submission were at least aware of the previous publication, unless they convince us otherwise.

John T. Cathey
Senior Editor

Mahmoud Aljurf, MD
Editor-in-Chief

Reply

Dear Dr. Afifi: Thank you for raising an important issue and providing the data on duplicate publication in the region. In response to your comment about our policy, our decision to ban all authors on manuscripts submitted to the Annals that are found to have been previously published is based on an assumption that all authors partook in the study and were aware of the original publication and the second submis-