Plagiarism in the Context of Education and Evolving Detection Strategies

Armen Yuri Gasparyan,[1]
Bekaidar Nurmashev,[2]
Bakhytzhan Seksenbayev,[2]
Vladimir I. Trukhachev,[3]
Elena I. Kostyukova,[4]
and George D. Kitas[1,5]

[1]Departments of Rheumatology and Research and Development, Dudley Group NHS Foundation Trust (Teaching Trust of the University of Birmingham, UK), Russells Hall Hospital, Dudley, West Midlands, UK; [2]South Kazakhstan State Pharmaceutical Academy, Shymkent, Kazakhstan; [3]Stavropol State Agrarian University, Stavropol, Russian Federation; [4]Department of Accounting Management, Stavropol State Agrarian University, Stavropol, Russian Federation; [5]Arthritis Research UK Epidemiology Unit, University of Manchester, Manchester, UK

Address for Correspondence:
Armen Yuri Gasparyan, MD
Departments of Rheumatology and Research and Development, Dudley Group NHS Foundation Trust (Teaching Trust of the University of Birmingham, UK), Russells Hall Hospital, Pensnett Road, Dudley DY1 2HQ, West Midlands, UK
E-mail: a.gasparyan@gmail.com

Plagiarism may take place in any scientific journals despite currently employed anti-plagiarism tools. The absence of widely acceptable definitions of research misconduct and reliance solely on similarity checks do not allow journal editors to prevent most complex cases of recycling of scientific information and wasteful, or 'predatory,' publishing. This article analyses Scopus-based publication activity and evidence on poor writing, lack of related training, emerging anti-plagiarism strategies, and new forms of massive wasting of resources by publishing largely recycled items, which evade the 'red flags' of similarity checks. In some non-Anglophone countries 'copy-and-paste' writing still plagues pre- and postgraduate education. Poor research management, absence of courses on publication ethics, and limited access to quality sources confound plagiarism as a cross-cultural and multidisciplinary phenomenon. Over the past decade, the advent of anti-plagiarism software checks has helped uncover elementary forms of textual recycling across journals. But such a tool alone proves inefficient for preventing complex forms of plagiarism. Recent mass retractions of plagiarized articles by reputable open-access journals point to critical deficiencies of current anti-plagiarism software that do not recognize manipulative paraphrasing and editing. Manipulative editing also finds its way to predatory journals, ignoring the adherence to publication ethics and accommodating nonsense plagiarized items. The evolving preventive strategies are increasingly relying on intelligent (semantic) digital technologies, comprehensively evaluating texts, keywords, graphics, and reference lists. It is the right time to enforce adherence to global editorial guidance and implement a comprehensive anti-plagiarism strategy by helping all stakeholders of scholarly communication.

Keywords: Plagiarism; Information Retrieval; Similarity Detection; Research Activity; Retraction of Publication as Topic; Research and Development; Publication Ethics

INTRODUCTION

Scholarly publications are the essence of scientific research, networking, and exploration of new areas of knowledge. Reporting original data and formulating new ideas in one's own words are traditionally viewed as the foundation of effective publishing and post-publication communication.

Plagiarism is an act of misappropriation of others' intellectual property, including but not limited to scholarly texts, research methods, graphics, and ideas (1). Along with fabrication and falsification, it is classed as research misconduct. The etymology of the term "plagiarism," which means literary piracy, refers to the Latin words plagiārius ("kidnapper") and plaga ("hunting net"), stemming from the Ancient Greek root word πλέκειν (ple-kein, "to weave," "to braid"). The modern Greek term λογοκλοπία (logoklopía, from λογο- [logo-, "word"] and κλέπτω [klépto, "steal"]) points to the wrongful appropriation of words.

Failure to obtain permission to reproduce previously published material and to acknowledge primary sources are the main components of the misappropriation (2). Although substantial un-attributed textual copying may lead to authorship disputes and copyright infringement with legal consequences (3,4), plagiarism is widely viewed as a distinct ethical issue, necessitating rewriting, rejection, or retraction of copied texts and whole papers, public shaming and expulsion of plagiarists with their temporary or permanent barring from publishing. The seriousness of the charge and severity of the penalty depend on the specific motives of plagiarists (intentional, unintentional, or accidental theft), their awareness of related ethical norms, language proficiency, context, and volume of copying (5-8). Publication experts judge intentional and unintentional research misconduct differently. Despite the fact that both forms of misconduct distort the scientific record and should not be tolerated (9), there are some who argue that unintentional misconduct is "less serious" and "less harmful" (10).

Debates also surround self-plagiarism, or text recycling, which

is viewed by some authors as an acceptable form of reusing own words and writings. This is a vexing problem for those who work in a single field and publish articles with overlapping definitions, descriptions of methods, and references to own works. It is also a problem for academic institutions, where students' dissertations containing their own published articles as whole chapters are viewed as self-plagiarized. In cases of reusing copyright-protected material, self-plagiarists may even face legal actions by primary publishers. To avoid ethical and legal issues, authors are advised to avoid adapting own previous publications and make an extra effort when (re)writing new texts (11,12).

Plagiarized items may appear in any journals, regardless of their scope, prestige of the publisher, geographical location, indexing status, citation impact, rigor of peer review, and formal adherence to ethical guidelines. However, indexed sources with low reader attention and minimal citations are more likely to accommodate illegitimately copied material (13). In fact, soft or non-existent peer review and essential editorial checks in 'predatory' journals attract inexperienced and/or dishonest authors, who may submit redundant and plagiarized manuscripts. Such journals exploit the gold open-access publishing model and charge their authors without providing quality services for distributing validated and innovative information (14). There are no exact estimates of the prevalence of plagiarism in scholarly publications due to uncertainties around its definition and deficiencies of current detection strategies. Analysis of corrective measures across indexed journals can help explore trends in unethical copying. A cross-sectional study of 134 retractions by BioMed Central journals in the period 2000 to 2015 identified textual plagiarism as a common reason of literature correction (16%) (15). Another analysis of more than 2,000 retracted PubMed-indexed articles revealed that 9.8% of the retractions were due to plagiarism, which became a frequent form of research misconduct since 2005 (16). Based on this study, Japan, China, India, Korea, Italy, Turkey, Iran, and France were identified as countries with a relatively high incidence of plagiarism. Finally, an analysis of 835 retracted papers indexed in PubMed between 2008 and 2012, demonstrated high rates of plagiarism-related retraction among authors from Italy (16 out of 24 retractions, 66.7%), India (18 out of 49, 36.7%), and China (24 out of 143, 16.8%) (17).

The reasons and forms of the misconduct are complex and poorly explored in the context of various linguistic and professional backgrounds. Poor writing, paraphrasing, and referencing skills are common and easily identifiable reasons of misappropriation of others' intellectual property, warranting advanced academic English writing courses for researchers and authors from non-Anglophone countries (18). There are still many instances of plagiarism in non-English publications, which are often overlooked and not taken seriously by research managers and editors' associations. Importantly, the growing demand for

publications, a major criterion for academic promotion, has created an impetus for commercial editing agencies in some countries to sell apparently copied manuscripts to inexperienced and careless authors (19).

Despite the fact that the intellectual theft is not a new phenomenon, its implications have become dramatic in these times of expanding online databases and archiving scientific evidence (20). Plagiarism is now viewed as a major threat to scientific evidence accumulation, which relies entirely on originality and transparency of scholarly publications (21). Unwitting authors, who perform systematic searches and retrieve literature from evidence-based databases, may cite such unethical items before and after their retractions, further damaging the credibility of current standards of research reporting.

With the constantly expanding number of indexed journals, the likelihood of encountering recycled, redundant, or otherwise unethical items is increasing. In the era of digital technologies and prospering commercial editing services, editors and readers encounter ever more cases of plagiarism, which pass unnoticed at checks by anti-plagiarism software.

PUBLICATION ACTIVITY

The dynamics and patterns of global interest to the issue can be explored by a snapshot analysis of searches through Scopus. As of March 31, 2017, there are 4,924 items tagged with the term "Plagiarism" in their titles, abstracts, or keywords, with date range of 1831 to 2017 (Fig. 1). The first 2 indexed letters of charges of plagiarism were published in *The Lancet* back in 1831. For the following 150 years, annual publication activity was negligible. In 1980, 17 tagged items were published, followed by 35 years of booming activity. Two peaks with 427 and 457 indexed documents are recorded in 2013 and 2016, respectively. The latest
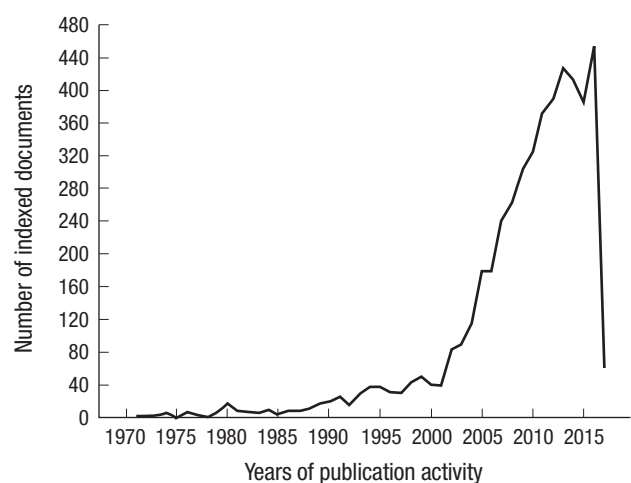

Fig. 1. Number of Scopus-indexed items tagged with the term "Plagiarism" in 1970–2017 (as of March 31, 2017).

article covers the issue of plagiarism of ideas, which can be traced by semantic analysis (22).

The top 5 sources in the field are *Lecture Notes in Computer Science* (n = 121), *Nature* (n = 104), *Notes and Queries* (n = 84), *Science* (n = 48), and *Science and Engineering Ethics* (n = 46). The top 5 prolific authors are Paolo Rosso (Valencia, Spain, number of tagged publications = 47), Benno Maria Stein (Weimar, Germany, n = 24), Miguel Roig (Queens, NY, USA, n = 22), Martin Potthast (Weimar, Germany, n = 20), Naomie Binti Salim (Skudai, Malaysia, n = 19). The main area of professional interest of these top 5 authors is computer science, except for Miguel Roig (psychology). The USA is the leading country in the field with 1,218 documents, followed by the UK (n = 469), India (n = 238), Australia (n = 212), and China (n = 197). The largest proportion of tagged documents are categorized as articles (n = 1,923, 39.1%), followed by conference papers (n = 1,011, 20.5%) and editorials (n = 527, 10.7%). Remarkably, there are 81 (1.6%) documents categorized as errata, including 29 items related to Iran and 53 items in the field of medicine. The majority of these documents (n = 45) were published in 2016 to notify about retractions of predominantly Iranian articles (n = 28), and specifically 21 items from *Diagnostic Pathology* (BioMed Central).

The leading subject areas are social sciences (n = 1,647, 33.4%), computer science (n = 1,290, 26.2%) and medicine (n = 1,235, 25.1%). The majority of items are in English (n = 4,575), followed by those in Spanish (n = 92) and German (n = 78). The top 10 highly-cited items attracted from 1,642 to 190 citations, with the landmark study of reasons for retractions cited 239 times (16). Finally, there are 61 tagged items, which are cited at least 61 times (h-index = 61).

PLAGIARISM IN THE CONTEXT OF EDUCATION

The large number of tagged items in social sciences points to global problems in under- and postgraduate education confounding the lack of understanding of plagiarism, cheating, and other forms of academic dishonesty among students who may choose to write and publish articles later in life. Comparative analyses suggest that young, undergraduate students, those in business studies and engineering, and residents of non-Anglophone countries frequently cheat and violate established norms of publication ethics (23-25).

In some Asian countries, plagiarism is viewed as a social phenomenon rooted in the dogmatic system of education, encouraging reuse of textbook information and suppressing creative thinking and generation of untested ideas (26). Poor education and unethical source use lead to plagiaristic writing by Taiwanese college students (27). Similar conclusions were drawn based on anti-plagiarism software checks of Malaysian undergraduate students' essays, containing large chunks of copied texts from easily accessible online sources without proper paraphrasing

and referencing (28). Finally, a report from Korea found that first-year medical students, who are unaware of what constitutes research misconduct, often write their papers by copying material from a limited number of easily accessible online platforms, such as Google, and do not cite the sources used (29).

A cross-cultural study of plagiarism perceptions suggested that students from Germany are more sensitive toward plagiarism and better skilled to identify academic dishonesty than their Turkish and Georgian peers (30). Likewise, a survey of Australian and Chinese undergraduates concluded that Australians are more negative toward plagiarism (31). Although cultural, linguistic and psychological factors determine the behavior of plagiarists, the lack of institutional anti-plagiarism policies is believed to play a more important role (32). Academic institutions across the world differ widely in their definitions of plagiarism, practices of preventing academic dishonesty, and research methodology courses for students, deficiencies of which lead to instances of unintentional plagiarism in some countries (33-35).

The lack of undergraduate courses on plagiarism is reflected in poor citing and referencing skills of Iranian medical students, who may plagiarize in their first articles but become more conversant with research integrity by publishing more (36). Similar trends are observed elsewhere in the world, suggesting that technological advances and educational initiatives reduce the occurrence of 'copy-and-paste writing' over time (37).

A large survey of Pakistani Bachelor degree medical students (n = 421) proved that training on research ethics improve their perception of plagiarism to a level comparable to that of the faculty (38). For non-medical international Master students, librarian-guided courses on legal and ethical aspects of research, proper citing, referencing and paraphrasing have also proved successful for better understanding of plagiarism and its consequences (39).

Oversights in the system of education along with numerous short-cuts in the process of preparing student works add to the problem of plagiarism. Students with poor time management, inadequate English writing skills and lacking sufficient support by their mentors often refer to commercial editing agencies for ghost-writing or otherwise unethical services (40,41). The globally mushrooming 'contract cheating' services are also good examples of how outsourced writing assignments turn into fraud and plagiarism (42). The term was first coined by experts in computer science from Birmingham City University (Birmingham, UK), who described unethical online bid requests to get students' coursework completed by third parties (43). Students committing such an academic dishonesty usually lack English language skills and seek advanced professional services by paid contractors (44).

DETECTION OF PLAGIARISM

A large proportion of Scopus-indexed items on plagiarism relates to computer science (26.2%), reflecting the focus of specialists on digital systems for tracking overlaps in scholarly articles. Over the past few decades, reuse of words and acceptable thresholds of textual matching have been thoroughly explored in the context of manual and software checks (45). Reuse of words is an elementary form of copying, which can be easily tracked either by manual searches through Google or by software-assisted tests, provided primary sources of interest are covered by related platforms. Non-Anglophone and novice native English-speaking authors alike are now advised to check their manuscripts by advanced text-matching software, such as iThenticate®, to detect and rewrite copied parts, and avoid accusations of misconduct (46). In an attempt to upgrade their journal ethics, editors are increasingly employing powerful anti-plagiarism software for checks at various stages of the manuscript processing (47-49). *The Lancet* journals, for example, perform anti-plagiarism checks with iThenticate® for all review manuscripts to avoid substantial textual overlaps with already published material (50).

Remarkably, the iThenticate® database currently has access to nearly 226,000 journals, and more than 1,300 publishers worldwide employ its similarity check services (51). While this powerful anti-plagiarism software with access to massive volumes of scholarly sources is now widely employed by leading publishers, including Elsevier, the Nature Publishing Group, and Springer, it is still unaffordable for start-up and small publishers and standalone journals with limited funding. Established publishers started trialling the software back in June 2008, and fees at that time started out at $0.75 per article (52). However, a more recent report indicates that the iThenticate® screening of all 300–350 annual submissions would cost the *American Journal of Neuroradiology*, the official organ of the American Society of Neuroradiology, approximately $6,800 (53).

The relatively high cost of the similarity checks is a barrier for testing all submissions by publishers in developing countries. Testing only suspicious manuscripts, and particularly reviews, is perhaps justifiable for these publishers, but such a strategy opens the gates for numerous research papers with inappropriately (copy-and-paste) written introduction and discussion sections. It is also known that software is powered to detect copied texts, but fails to highlight parts extensively edited by commercial agents, recycling both words and ideas and fooling the anti-plagiarism system (54). The practice of copying texts and substituting separate words with synonyms to evade plagiarism detection by software is known as rogeting (after Roget's thesaurus). An appalling example of rogeting was described in connection with a predatory journal, where a book chapter was entirely copied, edited by using synonyms, and re-published (55).

Scientists continue exploring other automatic options for detecting similarities beyond identical textual matches, which may help detect sophisticated forms of plagiarism. New intelligent computer programs are developed to track inappropriate paraphrasing and summarizing, manipulation of words and synonyms, cross-language plagiarism, and incorrect copying of references and images. There are numerous paraphrasing tricks, including copying verbatim and substituting separate words, which are often used in combination to mask text-borrowing and self-plagiarism. Such writing strategies are known to linguists but not easily recognized by non-experts (56). Fortunately, a computer-based analysis of lexical, syntactic, and semantic features is now implemented to track inappropriate paraphrasing in suspicious sentences and larger passages of texts (57,58). Promising are also results of studies of cross-language text-borrowing by non-native English speakers, who may copy, translate to English, and edit texts published in their mother tongue (59,60). Intelligent systems are mainly designed to recognize incorrect processing of German, French, and other Roman texts, though few recent studies have also focused on other languages (61). One of the proposed principles of unmasking cross-language plagiarism is based on analysing keywords in suspicious texts (62).

Citation-based plagiarism detection is another evolving concept that relies on similarities of order and proximity of references in texts (63). Related methods employ analyses of published reference lists visible in large bibliographic databases such as Scopus and Web of Science. Combining text and reference list checks increases the chances of detecting plagiarism (64).

There are also basic search services, such as Google Images, which allow users to retrieve already published graphical material by entering keywords related to the image file name and adjacent text. More advanced algorithms are now designed to track image misappropriation by evaluating the graphical structure and quantifying similarities between original and suspicious files (65).

Advances in the visualization of scholarly works, including their exposure to social networking sites, may improve the efficiency of detecting and quantifying plagiarism by emerging electronic gadgets (66). Ultimately, it is believed that a universal plagiarism detection system of cross-checking submissions to all peer-reviewed journals can minimize the number of retractions due to plagiarism, duplication and simultaneous submissions (49).

LESSONS LEARNED FROM RETRACTIONS

Different types of plagiarism are currently reported in retraction notices, ranging from copy-and-paste writing to more sophisticated manipulations. Examining the motivation and con-

text in such cases is critical for preventing misconduct in the future. Although the number of retractions due to plagiarism is not high (81 [1.6%] of documents in Scopus), the high prevalence of such cases in the medical literature, primarily affecting rapidly developing disciplines and emerging scientific powers, is worrying (67).

In November 2016, retractions of 58 Iranian articles by impact-factor journals of Springer and BioMed Central have been widely discussed in social media (68), conveying a strong message to all stakeholders of science communication. The most affected journals were *Tumor Biology* (25 retractions) and *Diagnostic Pathology* (23 retractions). The retracted articles were authored by 282 early-career and experienced researchers, who covered issues in cancer, veterinary science, pharmacology, and parasitology and some of whom previously published in predatory journals. An investigation, which was initially prompted by a whistleblower, found evidence of plagiarism in 93% of BMC and 70% of Springer articles that was not detected by software because of manipulations at the submission and peer review (69). Suspicions were also raised that the retracted articles could be produced by a random paper generator program, which were not confirmed. As a result of the retractions, serial plagiarists were barred from publishing, instructions of the journals were amended to exclude manipulations with authorship and peer review but no changes were announced to upgrade plagiarism detection. That case was followed by 5 more retractions of largely plagiarized Chinese papers on cancer by a Brazilian open-access journal *Genetics and Molecular Research* (70).

CONCLUSION

Plagiarism is a plague of current science and a sign of pseudoscience. Its detection and prevention strategies require a comprehensive approach by all stakeholders of science communication, based on their awareness of the global and local trends in misappropriation of intellectual property. Relying on any single anti-plagiarism tool is not effective. Even powered anti-plagiarism software, designed to track textual overlaps and calculate similarity scores, can be fooled by plagiarists, resorting on manipulative paraphrasing and editing techniques. Editors should manually check all manuscripts after receiving the software-generated similarity percentages to avoid false negative and positive results (71).

Among all preventive anti-plagiarism strategies, those relying on human factor are probably more reliable. A more active role of authors, reviewers, editors, and readers in the continuum of prevention is advisable. Inexperienced authors need to be updated by research supervisors on what constitutes plagiarism of words, ideas, and graphics. Their disclaimers about honest writing, which are currently requested by some journals (e.g., *Journal of Pakistan Medical Association*), may minimize the chanc-

es of and increase accountability for any wrongdoing. Reviewers and readers, or whistleblowers, may recognize sentences, larger passages of texts and order of references recycled from their own works and inform editors. Editors, in turn, may prevent certain types of plagiarism by carefully selecting reviewers and performing additional checks of texts and data. They should also be alerted of plagiarism by some reviewers, who may steel data and ideas from processed manuscripts and publish their own unethical articles (72). Vigilance is also required when students' works are processed as some of them may misappropriate unpublished raw data and publish papers without knowledge of their supervisors and owners of the data (73).

Plagiarism takes different forms and affects various indexed and non-indexed sources in any language. However, this misconduct is relatively easy to uncover in widely visible English sources. As a prime example, even top scholarly journals have suffered from misappropriation of words and sentences by non-Anglophone and inexperienced authors (6,74). To a certain extent, such a misconduct does not affect the validity of research reports and is correctable (8). Self-correction in top journals can limit and prevent the growth of unethical papers (75). What is more pressing is that most non-English and non-indexed low-profile periodicals, covering research of non-mainstream science authors, remain largely in the shadow and escape attention of eagle-eyed readers.

In recent years, predatory publishing has emerged as a serious threat to non-mainstream science (76). The absence of anti-plagiarism policies in such journals leads to massive online publishing of entirely copied and unchecked papers (77). The scale of the misconduct and carelessness of the authors may point to the involvement of paper-generating machines and editing agencies with commercial interests and ignorance of norms of scholarly publishing. Recent cases of mass retractions may also be associated with indiscriminate targeting of predatory editing agencies and individual authors, expanding their activities and exploiting the deficiencies of the similarity checks by indexed journals.

Intelligent (semantic) technologies may help detect more cases of plagiarism worldwide and across academic disciplines with inevitable retractions and dire consequences for plagiarists (78-81). Authors and editors may curb the problem by adhering to global editorial recommendations, such as those by the Committee on Publication Ethics (COPE) and the International Committee of Medical Journal Editors (ICMJE). COPE is the largest community of editors with more than 10,000 members from diverse professional and language backgrounds, who are offered a forum for discussing and resolving complex issues of plagiarism (82). ICMJE with its 15 members, including representatives of top-tier general medical journals, offer annually updated recommendations on proper writing, research reporting, targeting ethical journals, and avoiding research miscon-

duct (83). The ever-increasing support of the global editorial associations is instrumental for improving writing, reviewing and editing skills of all stakeholders of science communication. Consequently, poor writing and (un)intentional plagiarism is no longer an excuse for wasting resources of the scholarly community (84,85).

DISCLOSURE

The authors have no potential conflicts of interest to disclose.

AUTHOR CONTRIBUTION

Conceptualization: Gasparyan AY, Nurmashev B, Kostyukova EI, Kitas GD. Methodology: Gasparyan AY, Nurmashev B, Seksenbayev B. Writing - original draft: Gasparyan AY, Nurmashev B. Writing - review & editing: Gasparyan AY, Nurmashev B, Seksenbayev B, Kostyukova EI, Kitas GD.

ORCID

Armen Yuri Gasparyan https://orcid.org/0000-0001-8749-6018
Vladimir I. Trukhachev https://orcid.org/0000-0003-4676-5407
Elena I. Kostyukova https://orcid.org/0000-0002-8393-6001
George D. Kitas https://orcid.org/0000-0002-0828-6176

REFERENCES

1. Roig M. Avoiding unethical writing practices. *Food Chem Toxicol* 2012; 50: 3385-7.

2. Council of Science Editors (US). CSE's white paper on promoting integrity in scientific journal publications, 2012 update [Internet]. Available at http://cseditors.wpengine.com/wp-content/uploads/entire_whitepaper.pdf [accessed on 31 March 2017].

3. U.S. Department of Health and Human Services, Office of Research Integrity. ORI policy on plagiarism [Internet]. Available at https://ori.hhs.gov/ori-policy-plagiarism [accessed on 31 March 2017].

4. U.S. Department of Health and Human Services, Office of Research Integrity. Copyright infringement, fair use, and plagiarism [Internet]. Available at https://ori.hhs.gov/plagiarism-18 [accessed on 31 March 2017].

5. Wittmaack K. Penalties plus high-quality review to fight plagiarism. *Nature* 2005; 436: 24.

6. Roig M. Plagiarism: consider the context. *Science* 2009; 325: 813-4.

7. Resnik DB, Dinse GE. Scientific retractions and corrections related to misconduct findings. *J Med Ethics* 2013; 39: 46-50.

8. Chaddah P. Not all plagiarism requires a retraction. *Nature* 2014; 511: 127.

9. Self-plagiarism: unintentional, harmless, or fraud? *Lancet* 2009; 374: 664.

10. Helgesson G, Eriksson S. Plagiarism in research. *Med Health Care Philos* 2015; 18: 91-101.

11. Supak-Smočić V, Bilić-Zulle L. How do we handle self-plagiarism in submitted manuscripts? *Biochem Med (Zagreb)* 2013; 23: 150-3.

12. Stone TE, Conway J. Editorial: Self-plagiarism prevention and management at Nursing & Health Sciences. *Nurs Health Sci* 2017; 19: 1-4.

13. de Almeida RM, Catelani F, Fontes-Pereira AJ, Gave Nde S. Retractions in general and internal medicine in a high-profile scientific indexing database. *Sao Paulo Med J* 2016; 134: 74-8.

14. Beall J. Dangerous predatory publishers threaten medical research. *J Korean Med Sci* 2016; 31: 1511-3.

15. Moylan EC, Kowalczuk MK. Why articles are retracted: a retrospective cross-sectional study of retraction notices at BioMed Central. *BMJ Open* 2016; 6: e012047.

16. Fang FC, Steen RG, Casadevall A. Misconduct accounts for the majority of retracted scientific publications. *Proc Natl Acad Sci U S A* 2012; 109: 17028-33.

17. Amos KA. The ethics of scholarly publishing: exploring differences in plagiarism and duplicate publication across nations. *J Med Libr Assoc* 2014; 102: 87-91.

18. Heitman E, Litewka S. International perspectives on plagiarism and considerations for teaching international trainees. *Urol Oncol* 2011; 29: 104-8.

19. Shahghasemi E, Akhavan M. Confessions of academic ghost authors: the Iranian experience. *Sage Open* 2015; 5: 1-7.

20. de Vasconcelos SM, Roig M. Prior publication and redundancy in contemporary science: are authors and editors at the crossroads? *Sci Eng Ethics* 2015; 21: 1367-78.

21. Rathod SD. Combating plagiarism: a shared responsibility. *Indian J Med Ethics* 2010; 7: 173-5.

22. Vani K, Gupta D. Detection of idea plagiarism using syntax–semantic concept extractions with genetic algorithm. *Expert Syst Appl* 2017; 73: 11-26.

23. Park C. In other (people's) words: plagiarism by university students--literature and lessons. *Assess Eval High Educ* 2003; 28: 471-88.

24. Wheeler G. Plagiarism in the Japanese universities: truly a cultural matter? *J Second Lang Writ* 2009; 18: 17-29.

25. Arda B. Publication ethics from the perspective of PhD students of health sciences: a limited experience. *Sci Eng Ethics* 2012; 18: 213-22.

26. Chaurasia A. Stop teaching Indians to copy and paste. *Nature* 2016; 534: 591.

27. Chien SC. Taiwanese college students' perceptions of plagiarism: cultural and educational considerations. *Ethics Behav* 2017; 27: 118-39.

28. Zangenehmadar S, Hoon TB. ESL undergraduates' patterns of plagiarism in academic essays based on print or internet sources. *Pertanika J Soc Sci Humanit* 2017; 25: 205-20.

29. Kim KJ, Hwang JY, Lee DW, Shim MS. Medical student plagiarism in problem-based learning courses. *Med Educ Online* 2016; 21: 30537.

30. Kayaoğlu MN, Erbay Ş, Flitner C, Saltaş D. Examining students' perceptions of plagiarism: a cross-cultural study at tertiary level. *J Furth High Educ* 2016; 40: 682-705.

31. Ehrich J, Howard SJ, Mu C, Bokosmaty S. A comparison of Chinese and Australian university students' attitudes towards plagiarism. *Stud High Educ* 2016; 41: 231-46.

32. Hu G, Sun X. Institutional policies on plagiarism: the case of eight Chinese universities of foreign languages/international studies. *System* 2017; 66: 56-68.

33. Bretag T. Challenges in addressing plagiarism in education. *PLoS Med* 2013; 10: e1001574.

34. Kokkinaki AI, Demoliou C, Iakovidou M. Students' perceptions of plagiarism and relevant policies in Cyprus. *Int J Edu Integr* 2015; 11: 3.

35. Varghese J, Jacob M. Do medical students require education on issues re-

lated to plagiarism? *Indian J Med Ethics* 2015; 12: 82-7.

36. Gharedaghi MH, Nourijelyani K, Salehi Sadaghiani M, Yousefzadeh-Fard Y, Gharedaghi A, Javadian P, Morteza A, Andrabi Y, Nedjat S. Knowledge of medical students of Tehran University of Medical Sciences regarding plagiarism. *Acta Med Iran* 2013; 51: 418-24.

37. Curtis GJ, Vardanega L. Is plagiarism changing over time? A 10-year time-lag study with three points of measurement. *High Educ Res Dev* 2016; 35: 1167-79.

38. Rathore FA, Waqas A, Zia AM, Mavrinac M, Farooq F. Exploring the attitudes of medical faculty members and students in Pakistan towards plagiarism: a cross sectional survey. *PeerJ* 2015; 3: e1031.

39. Gunnarsson J, Kulesza WJ, Pettersson A. Teaching international students how to avoid plagiarism: librarians and faculty in collaboration. *J Acad Librariansh* 2014; 40: 413-7.

40. Shaw C. International students are turning to proofreading agencies to get support [Inrernet]. Available at https://www.theguardian.com/higher-education-network/blog/2014/apr/09/international-students-proofreading-academic-writing-support [accessed on 31 March 2017].

41. Jones-Berry S. Online market for nursing essays bigger than ever. *Nurs Stand* 2016; 31: 12-3.

42. Draper MJ, Ibezim V, Newton PM. Are Essay Mills committing fraud? An analysis of their behaviours vs the 2006 Fraud Act (UK). *Int J Educ Integr* 2017; 13: 3.

43. Lancaster T, Clarke R. The phenomena of contract cheating. In: Roberts TS, editor. Student Plagiarism in an Online World: Problems and Solutions. Hershey, PA, Information Science Reference, 2008, p144-59.

44. Rigby D, Burton M, Balcombe K, Bateman I, Mulatu A. Contract cheating & the market in essays. *J Econ Behav Organ* 2015; 111: 23-37.

45. Rathore FA, Farooq F. Plagiarism detection softwares: useful tools for medical writers and editors. *J Pak Med Assoc* 2014; 64: 1329-30.

46. Li Y. Text-based plagiarism in scientific publishing: issues, developments and education. *Sci Eng Ethics* 2013; 19: 1241-54.

47. Supak-Smolcić V, Simundić AM. Biochemia Medica has started using the CrossCheck plagiarism detection software powered by iThenticate. *Biochem Med (Zagreb)* 2013; 23: 139-40.

48. Thordarson DB. Plagiarism and journal ethics. *Foot Ankle Int* 2013; 34: 1611.

49. Jarić I. High time for a common plagiarism detection system. *Scientometrics* 2016; 106: 457-9.

50. Kleinert S; on behalf of the editors of all Lancet journals. Checking for plagiarism, duplicate publication, and text recycling. *Lancet* 2011; 377: 281-2.

51. iThenticate (US). Crossref similarity check powered by iThenticate [Internet]. Available at http://www.ithenticate.com/products/crosscheck [accessed on 31 March 2017].

52. Butler D. Journals step up plagiarism policing. *Nature* 2010; 466: 167.

53. Kalnins AU, Halm K, Castillo M. Screening for self-plagiarism in a subspecialty-versus-general imaging journal using iThenticate. *AJNR Am J Neuroradiol* 2015; 36: 1034-8.

54. Chaddah P. Pursuing knowledge creation, India needs a policy on 'plagiarism cells.' *Curr Sci* 2014; 106: 349.

55. Neuroskeptic (US). Plagiarism: copy, paste, thesaurus? [Internet]. Available at http://blogs.discovermagazine.com/neuroskeptic/2015/02/07/plagiarism-thesaurus/#.WQXZlBFOXcs [accessed on 31 March 2017].

56. Sun YC, Yang FY. Uncovering published authors' text-borrowing practices: paraphrasing strategies, sources, and self-plagiarism. *J Engl Acad Pur-*

poses 2015; 20: 224-36.

57. Chitra A, Rajkumar A. Plagiarism detection using machine learning-based paraphrase recognizer. *J Intell Syst* 2016; 25: 351-9.

58. Hussein AM. Paraphrasing: solution to the problem of plagiarism. *Asian J Inf Technol* 2016; 15: 2743-6.

59. Franco-Salvador M, Gupta P, Rosso P, Banchs RE. Cross-language plagiarism detection over continuous-space- and knowledge graph-based representations of language. *Knowl Base Syst* 2016; 111: 87-99.

60. Franco-Salvador M, Rosso P, Montes-y-Gómez M. A systematic study of knowledge graph analysis for cross-language plagiarism detection. *Inf Process Manage* 2016; 52: 550-70.

61. Asghari H, Mohtaj S, Fatemi O, Faili H, Rosso P, Potthast M. Algorithms and corpora for Persian plagiarism detection. *CEUR Workshop Proc* 2016; 1737: 135-44.

62. Ehsan N, Shakery A. Candidate document retrieval for cross-lingual plagiarism detection using two-level proximity information. *Inf Process Manag* 2016; 52: 1004-17.

63. Gipp B, Meuschke N, Breitinger C. Citation-based plagiarism detection: practicability on a large-scale scientific corpus. *J Assoc Inf Sci Technol* 2014; 65: 1527-40.

64. Pertile SD, Moreira VP, Rosso P. Comparing and combining content- and citation-based approaches for plagiarism detection. *J Assoc Inf Sci Technol* 2016; 67: 2511-26.

65. Rabiu I, Salim N. Textual and structural approaches to detecting figure plagiarism in scientific publications. *J Theor Appl Inf Technol* 2014; 70: 356-71.

66. Zrnec A, Lavbič D. Social network aided plagiarism detection. *Br J Educ Technol* 2017; 48: 113-28.

67. Almeida RM, de Albuquerque Rocha K, Catelani F, Fontes-Pereira AJ, Vasconcelos SM. Plagiarism allegations account for most retractions in major Latin American/Caribbean databases. *Sci Eng Ethics* 2016; 22: 1447-56.

68. Callaway E. Publisher pulls 58 articles by Iranian scientists over authorship manipulation [Internet]. Avalable at http://www.nature.com/news/publisher-pulls-58-articles-by-iranian-scientists-over-authorship-manipulation-1.20916 [accessed on 31 March 2017].

69. Retraction Watch (US). Springer, BMC retracting nearly 60 papers for fake reviews and other issues [Internet]. Available at http://retractionwatch.com/2016/11/01/springer-bmc-retracting-nearly-60-papers-for-fake-reviews-and-other-issues/#more-45774 [accessed on 31 March 2017].

70. A M Duarte F. RETRACTION of articles with common plagiarism. *Genet Mol Res* 2016; 15: gmr1504005.

71. Lykkesfeldt J. Strategies for using plagiarism software in the screening of incoming journal manuscripts: recommendations based on a recent literature survey. *Basic Clin Pharmacol Toxicol* 2016; 119: 161-4.

72. Dyer O. Peer reviewer stole article and published it as his own. *BMJ* 2016; 355: i6768.

73. Yahaghi J, Beddu SB, Muda ZC. Plagiarism in publications using the unpublished raw data of archived research. *Sci Eng Ethics* 2017; 23: 635-6.

74. Yilmaz I. Plagiarism? No, we're just borrowing better English. *Nature* 2007; 449: 658.

75. Gasparyan AY, Ayvazyan L, Akazhanov NA, Kitas GD. Self-correction in biomedical publications and the scientific impact. *Croat Med J* 2014; 55: 61-72.

76. Gasparyan AY, Nurmashev B, Udovik EE, Koroleva AM, Kitas GD. Preda-

tory publishing is a threat to non-mainstream science. *J Korean Med Sci* 2017; 32: 713-7.

77. Jansen PA, Forget PM. Predatory publishers and plagiarism prevention. *Science* 2012; 336: 1380.

78. Jargin SV. Pathology in the former Soviet Union: scientific misconduct and related phenomena. *Dermatol Pract Concept* 2011; 1: 75-81.

79. Yalcintas A, Wible JR. Scientific misconduct and research ethics in economics: an introduction. *Rev Soc Econ* 2016; 74: 1-6.

80. Hong ST. Plagiarism continues to affect scholarly journals. *J Korean Med Sci* 2017; 32: 183-5.

81. Rai R, Sabharwal S. Retracted publications in orthopaedics: prevalence, characteristics, and trends. *J Bone Joint Surg Am* 2017; 99: e44.

82. Pierson CA. Avoiding ethics pitfalls in publishing: a perspective from COPE. *Oral Dis* 2017; 23: 416-9.

83. International Committee of Medical Journal Editors (US). Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals [Internet]. Available at http://icmje.org/icmje-recommendations.pdf [accessed on 17 May 2017].

84. Glänzel W, Braun T, Schubert A, Zosimo-Landolfo G. Coping with copying. *Scientometrics* 2015; 102: 1-3.

85. Kojima T, Barron JP. Protect yourself from plagiarism. *Jpn J Gastroenterol Surg* 2016; 49: 1179-80.