Original Article

Relevancy of Free Online Software and Websites in Detection of Plagiarism

Himel Mondal, Shaikat Mondal[1]

Department of Physiology, MKCG Medical College, Ganjam, Odisha, [1]Department of Physiology, Medical College and Hospital, Kolkata, West Bengal, India

Abstract

Background and Aim: Plagiarism is disgraceful to authors and opposed in the scientific community. Before submitting articles to journals, authors should check whether the article is unintentionally plagiarized or not. However, online premium plagiarism checking service may not be accessible to many authors. For those authors, there are many websites which offer plagiarism detection service free of cost. The aim of the study was to assess the efficacy of those websites in the detection of plagiarism. **Methods:** A paragraph of text with 2 verbatim plagiarized sentences, 2 inadequately paraphrased sentences, and 2 perceived unique sentences was used as a tool. This paragraph was checked for plagiarism in 23 popular websites. Results showed by websites were recorded and analyzed. Chi-square test with $\alpha = 0.05$ was used to compare the proportion of websites. **Results:** Among 23 websites, only 4 websites (17.39%) detected plagiarism above the expected level. Thirteen (56.52%) websites showed result of 0% plagiarism for the paragraph. Verbatim plagiarism was detected by 6 websites (26.09%) where only 1 website could detect inadequate paraphrasing. **Conclusion:** The majority of websites offering free plagiarism checking service failed to detect plagiarized content effectively. The level of detected plagiarism differs in the different website for same content. Hence, checking the manuscript in multiple websites would provide relatively better screening result for plagiarism.

Keywords: Article screening, citation, duplicate content, online plagiarism, paraphrase

Received: 01st August, 2017; ***Revised:*** 12th September, 2017; ***Accepted:*** 20th September, 2017

Introduction

Plagiarism in any form is disgraceful for authors, and it is being opposed by the scientific community. However, it is still prevalent globally. Authors face rejection from the majority of journals if plagiarized content is presented.[1] In addition, legal issues may arise immediately or in future for the presentation of plagiarized content.[2] When authors write scientific articles, they usually review previous published papers and gained knowledge about recent updates.[3] If they think about incorporating that in their article, they usually paraphrase the content and cite the article properly. Paraphrasing may be influenced by style and phrase previously read.[4] At this point, authors may seek check for plagiarism of the article.

It is extremely difficult to check plagiarism for the content which is not available in the electronic database (i.e., in World Wide Web). Hence, no online tool can provide 100% accurate plagiarism check.[5] However, a fair percentage of plagiarism can be checked by searching duplicate content in World Wide Web. There are websites which offer paid service of detection of plagiarism for articles. The paid services may not be accessible to many authors. For the welfare of those authors, there are numbers of websites which provide free plagiarism detection service online.

No previous study was conducted to ascertain the applicability of those websites in the detection of plagiarism. With this background, the aim of this study was to assess the efficacy of those websites in detection of duplicate content.

Materials and Methods

According to the aim of the study, a cross-sectional study was conducted from February to March 2017. For the study, a

Access this article online

Quick Response Code:

Website: www.ijcep.org

DOI: 10.4103/ijcep.ijcep_38_17

Address for correspondence: Dr. Himel Mondal, Department of Physiology, MKCG Medical College, Ganjam - 760 004, Odisha, India.
E-mail: himelmkcg@gmail.com

How to cite this article: Mondal H, Mondal S. Relevancy of free online software and websites in detection of plagiarism. Int J Clin Exp Physiol 2017;4:139-41.

paragraph of text was designed with an aim to check the level of plagiarism of that particular paragraph in different websites.

Preparation of study tool

The paragraph of text was prepared in Microsoft Word® file with 6 sentences having 3 types of content.

First 2 sentences were copied from the conclusion of a published original article whose text is available in Hypertext Markup Language (HTML) format in journal website. The article was also available in Portable Document Format (PDF). The issue of the journal was online since Thursday, December 29, 2016.[6] The article is accessible freely and is distributed under Creative Commons Attribution-NonCommercial-ShareAlike 3.0 license. Two sentences were composed of 39 words and 234 characters.

Second 2 sentences were first copied word for word in the Microsoft word file from the conclusion of a published original article. This article was also available both in HTML and PDF format and distributed under the same license as the previous journal. The issue of the journal was online since Monday, February 06, 2017.[7] It contained 60 words and 352 characters. We paraphrased the sentence inadequately with intact sentence structure. Paraphrased 2 sentences contained 61 words and 356 characters.

Third 2 sentences were written by the second author with a perception that the sentences were unique. It was written with 38 words and 203 characters.

Hence, the final paragraph was of 138 words and 793 characters. Overall plagiarism of the prepared paragraph was expected >60%. Expected plagiarism score for the first 2 sentences was 100%, second 2 sentences was 80%, and third 2 sentences was 0%.

Free plagiarism detection software or website selection

Windows operating system-based personal computer, available in the Department of Physiology for educational purpose, connected to broadband internet connection with >1 mbps connection was used for browsing the websites. A phrase - "free plagiarism check" was searched in Google search engine in Mozilla Firefox web browser.[8,9] It showed about 42,00,000 results in 0.51 s. Search result of first 6 pages was considered, and a list of websites was prepared. Google showed some advertisement of websites which offer plagiarism detection service on the search result pages. We excluded those websites from the study. According to the aim of the study, any websites which do not offer free plagiarism checking service were excluded. However, those websites provide free service for limited times for new users were included in the list. In the majority of the websites, the maximum allowed character limit for a single test was 1000 words. In some of the websites, the minimum words were 1000 characters. These websites were excluded from the study. After exclusion, Twenty-three websites were in the final list.

Screening of plagiarized paragraph

On Saturday, March 4, 2017, from 10:00 to 17:00 h, level of duplicate content of the prepared paragraph was checked online. Each website was opened in Mozilla internet browser one at a time. Text was copied from the Microsoft word file to clipboard and pasted in the particular text box on the website page and duplicate content was checked. In some of the websites, there was an option for uploading the document file for checking plagiarism. However, it was avoided wherever possible. After completion of analysis by website, results that were shown free of cost were noted down.

Statistical analysis

Websites were divided into two groups (viz., <60% and >60%) according to the detected percentage of duplicate content. Websites which detected <60% plagiarism for the paragraph were again divided into three groups (viz., 0%, 1%–30%, and 31%–60%) according to the detected percentage of duplicate content for the paragraph. Chi-square test was used to compare the proportion of websites. Two-tailed $P < 0.05$ was considered statistically significant. Statistical analysis was carried out in GraphPad Prism version 6.01 for Windows (GraphPad Software, La Jolla California, USA).

Results

Among 23 websites, 13 (56.52%) websites showed 0% plagiarism for the paragraph which was actually >60% plagiarized. Number of websites according to detected level of plagiarism is shown in Table 1. The result of the Chi-square test showed that the difference in proportion of websites was not occurred by chance.

Among 23 websites, 17 websites only showed total percentage of uniqueness or total percentage of plagiarism. Rest 6 websites showed segmental plagiarism report.

Discussion

Result of this study showed that majority of the website that provides free plagiarism checking services failed to detect plagiarism effectively. Hence, if authors intend to screen their articles through these websites, they may get wrong assessment

Table 1: Number of websites ($n=23$) according to shown result of plagiarism of a paragraph which had >60% plagiarized content

	Detected plagiarism <60%				Detected plagiarism >60%	χ^2	*P* of Chi-square test
	0%	1%-30%	31%-60%	Total			
Number of website*	13	1	5	19	4	9.78	0.0018†

*Accessed on March 04, 2017 between 10:00 and 17:00 h, †Statistically significant *P* value of Chi-square test

report for their article. Among the websites that detected more than expected level of plagiarism, one website showed 100% plagiarism report. The report of this website was also not concordant with our assumption.

Those websites which provide plagiarism report in segments showed a better result than those which shows overall percentage only. Six websites (26.09%) showed result in segments. Among that 6 websites, all websites detected plagiarism for the verbatim plagiarized sentences. However, websites which detected verbatim plagiarism factually cannot help in plagiarism detection. If authors copy sentences as a whole and paste it in his article with quotation with proper citation, it is not considered plagiarized.[10] In that case, if authors check it online for plagiarism, result would give a 100% plagiarism report. Hence, merely detection of duplicate content is not same as detection of plagiarism. In addition, if authors use someone else's sentences as a whole intentionally without proper citation, they already know about it. Hence, checking it online is unnecessary.

Only one among six websites could detect plagiarism for inadequately paraphrased sentences. Hence, inadequate paraphrase my gets clear report from majority of websites.

Hence, unintentional plagiarism can be checked with the help of freely available service by the websites to some extent [Table 1]. However, it should be kept in mind that result reflected by those websites may not be accurate. Hence, it is suggested that authors adopt a multiple check of their content in different websites. This would provide a better assessment of the article.

Limitations of the study

In this study, according to our logistics, we only assessed the websites which offer free plagiarism checking service. A similar study with premium websites would reflect more generalized and comparative result.

Conclusion

Majority of websites offering free plagiarism checking service failed to detect plagiarism from a plagiarized paragraph. Level of detected plagiarism varies among websites for the same text. Verbatim plagiarism is detected more effectively than paraphrased content. Free of cost plagiarism detection at authors end in multiple websites may augment the result of detected plagiarism.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

1. Thrower P. Eight Reasons I Rejected Your Article. Available from: https://www.elsevier.com/connect/8-reasons-i-rejected-your-article. [Last accessed on 2017 Mar 06].
2. Plagiarism Law. Available from: https://www.checkforplagiarism.net/plagiarism-law. [Last accessed on 2017 Mar 06].
3. Hewitt M. How to Search and Critically Evaluate Research Literature. The NIHR RDS for the East Midlands/Yorkshire & the Humber; 2007. Available from: http://www.worcester.ac.uk/documents/6_Critically_evaluate_research_literature_2009v2.pdf. [Last accessed on 2017 Mar 06].
4. Summarizing, Paraphrasing, and Quoting. Available from: http://www.isites.harvard.edu/icb/icb.do?keyword=k70847&pageid=icb.page350378. [Last accessed on 2017 Mar 06].
5. The Limitation of Every Plagiarism Checker. Available from: https://www.plagiarismtoday.com/2011/12/07/the-limitation-of-every-plagiarism-checker/. [Last accessed 2017 on Mar 06].
6. Mondal H, Mondal S. Sample size calculation to data analysis of a correlation study in Microsoft Excel®: A hands-on guide with example. Int J Clin Exp Physiol 2016;3:180-9. Available from: http://www.ijcep.org/text.asp?2016/3/4/180/196896. [Last accessed on 2017 Mar 05].
7. Das D, Mondal H. Evaluation of cardiac autonomic function in overweight males: A cross-sectional study. Adv Hum Biol 2017;7:23-6. Available from: http://www.aihbonline.com/text.asp?2017/7/1/23/199532. [Last accessed on 2017 Mar 5].
8. Available from: https://www.google.co.in/. [Last accessed on 2017 Mar 04].
9. Available from: https://www.mozilla.org/en-US/firefox/new/. [Last accessed on 2017 Mar 04].
10. What Constitutes Plagiarism? Harvard Guide to Using Sources. A Publication of the Harvard College Writing Program. Available from: http://www.isites.harvard.edu/icb/icb.do?keyword=k70847&pageid=icb.page342054. [Last accessed on 2017 Mar 08].