Are paid tools worth the cost? A prospective cross-over study to find the right tool for plagiarism detection

Abhishek Anil [a,1], Aswini Saravanan [a,1], Surjit Singh [a,1],
Muhammad Aaqib Shamim [a], Krishna Tiwari [a], Hina Lal [a], Shanmugapriya Seshatri [a],
Simi Bridjit Gomaz [a], Thoyyib P. Karat [b], Pradeep Dwivedi [a], Shoban Babu Varthya [a],
Rimple Jeet Kaur [a], Prakasini Satapathy [c], Bijaya Kumar Padhi [d], Shilpa Gaidhane [e],
Manoj Patil [f], Mahalaqua Nazli Khatib [g,***], Joshuan J. Barboza [h,**], Ranjit Sah [i,j,k,*]

[a] Department of Pharmacology, All India Institute of Medical Sciences, Jodhpur, Rajasthan-342005, India
[b] Department of Dermatology, Venereology and Leprosy, All India Institute of Medical Sciences, Jodhpur, Rajasthan-342005, India
[c] Global Center for Evidence Synthesis, Chandigarh-160036, India
[d] Department of Community Medicine and School of Public Health, Postgraduate Institute of Medical Education and Research, Chandigarh - 160012, India
[e] One Health Centre (COHERD), Jawaharlal Nehru Medical College, Datta Meghe Institute of Higher Education, Wardha - 442001, India
[f] Division of Evidence Synthesis, School of Epidemiology and Public Health and Research, Jawaharlal Nehru Medical College, Datta Meghe Institute of Higher Education, Wardha - 442001, India
[g] Division of Evidence Synthesis, Global Consortium of Public Health and Research, Datta Meghe Institute of Higher Education and Research, Wardha - 442001, India
[h] Escuela de Medicina, Universidad Cesar Vallejo, Trujillo, 13007, Peru
[i] Tribhuvan University Teaching Hospital, Kathmandu - 46000, Nepal
[j] Department of Clinical Microbiology, DY Patil Medical College, Hospital and Research Centre, DY Patil Vidyapeeth, Pune - 411000, Maharashtra, India
[k] Department of Public Health Dentistry, Dr. D.Y. Patil Dental College and Hospital, Dr. D.Y. Patil Vidyapeeth, Pune - 411018, Maharashtra, India

ARTICLE INFO

ABSTRACT

Background: The increasing pressure to publish research has led to a rise in plagiarism incidents, creating a need for effective plagiarism detection software. The importance of this study lies in the high cost variation amongst the available options for plagiarism detection. By uncovering the advantages of these low-cost or free alternatives, researchers could access the appropriate tools

* Corresponding author. Infectious Diseases Fellowship, Clinical Research (Harvard Medical School), Global Clinical Scholars Research Training (Harvard Medical School), Systemic Review and Meta-analysis (London School of Hygiene and Tropical Medicine), Editorial Board Member: TMAID, ID Cases, BMC Infectious Diseases, Tribhuvan University Teaching Hospital, Kathmandu - 46000, Nepal.
** Corresponding author.
*** Corresponding author. Division of Evidence Synthesis, Global Consortium of Public Health and Research, Datta Meghe Institute of Higher Education, Wardha - 442001, India.
E-mail addresses: drabhishekanil@gmail.com (A. Anil), aswinisarvn29@gmail.com (A. Saravanan), sehmby_ss@yahoo.com (S. Singh), aaqibsh@gmail.com (M.A. Shamim), dockrishnatiwari@gmail.com (K. Tiwari), hina.k.lal@gmail.com (H. Lal), spriyasesshatri@gmail.com (S. Seshatri), simibridjitgomaz77@gmail.com (S.B. Gomaz), thoyyibpk@gmail.com (T.P. Karat), dr.prad99@gmail.com (P. Dwivedi), drshobanpgimer@gmail.com (S.B. Varthya), sidhurimple@yahoo.com (R.J. Kaur), contact@gces.network (P. Satapathy), bkpadhi@gmail.com (B.K. Padhi), drshilpagaidhane@gmail.com (S. Gaidhane), mpatil98dent@gmail.com (M. Patil), nazli.786@rediffmail.com (M.N. Khatib), jbarbozame@ucvvirtual.edu.pe (J.J. Barboza), ranjitsah@iom.edu.np (R. Sah).
[1] All three are first authors and equally contributed.

Small SEO tools
DupliChecker
Overall similarity index

for plagiarism detection. This is the first study to compare four plagiarism detection tools and assess factors impacting their effectiveness in identifying plagiarism in AI-generated articles.

Methodology: A prospective cross-over study was conducted with the primary objective to compare Overall Similarity Index(OSI) of four plagiarism detection software(iThenticate, Grammarly, Small SEO Tools, and DupliChecker) on AI-generated articles. ChatGPT was used to generate 100 articles, ten from each of ten general domains affecting various aspects of life. These were run through four software, recording the OSI. Flesch Reading Ease Score(FRES), Gunning Fog Index(GFI), and Flesch-Kincaid Grade Level(FKGL) were used to assess how factors, such as article length and language complexity, impact plagiarism detection.

Table 1
Basic characteristics of the four included software for plagiarism detection.

	iThenticate [8]	Grammarly [9]	Small SEO tools [10]	DupliChecker [11]
Plans	Only paid version (single, multiple, organization).	Only paid version (Grammarly Premium).	Free and paid plans (Weekly, Basic, Classic, Enterprise and Institute).	Free and paid premium plans.
Languages assessed	English, Spanish, Simplified Chinese, German, French, Russian, Portuguese, Japaunese, Italian, Dutch, Arabic, Polish, Turkish, Korean, Swedish.	Only English.	Supports 30+ languages like English, Deutsche, Espanol, Italiano, Korean, Francais, etc.	English, Spanish, Russian, Portuguese, Dutch, Indonesian, Italian, Arabic, Turkish, French, Svenska, Korean, Danish, Romanian, Norwegian, Vietnamese, Japanese and German.
Per search limit	**Single:** 25,000 words and below. **Multiple:** Up to 75,000 words. **For organization:** Customizable plans	In any 30-day period, upto 150,000 words. In any 24-h period, upto 50,000 words (17).	**Free:** Upto 1000 words **Weekly plan:** Upto 5000, **Basic:** Upto 10,000 words, **Classic**: Upto 20,000 words, **Enterprise and Institute**: Upto 30,000 words.	**Free plan:** Upto 1000 words **Premium plan:** Upto 25,000 words.
Other features of the software	Doc-to-Doc comparison	1. Spelling mistakes, confusing grammar issues, incorrect punctuation 2. Tone detector (analyzes your words, phrasing, and punctuation so that you can ensure your writing says what you mean) 3. Conciseness suggestions to tighten loose phrasing. 4. Restructuring lengthy sentences 5. Tone adjustment suggestions 6. Recommend inclusive language where it's appropriate 7. Informal to formal writing 8. Engagement suggestions (helps avoid repetitive patterns, replace dull vocabulary) 9. Fluency suggestions 10. Formatting suggestions 11. Highlight the passages that require citations and give you the resources you need to properly credit your sources.	Automatic rewriting tool, Keyword position, Domain authority checker, Backlink checker, etc.	Paraphrasing tool, reverse image search.
Cost of premium plans	**Single:** 100 dollars. Credits are valid for 12 months. Includes 5 free revisions. **Multiple:** 300 dollars. Credits are valid for 12 months. Includes 5 free revisions. **For organization:** Varies as per customization	**1. Individual plans** Annual: 11,799 Indian rupees Quarterly: 4899 Indian rupees Monthly: 2499 Indian rupees **2. Team plans** Grammarly business (minimum 3 members, billed annually): 15 USD/month/member.	Has a wide range of plans under basic, classic, enterprise and institute domains, with the cheapest being the weekly plan: 4.99 USD per week.	Has a wide range of weekly, monthly, and yearly plans. Starting from 5 USD per week.
Outcomes	Only the plagiarism.	Comments on performance score, word count, readability, vocabulary, etc. along with plagiarism percentage.	Percentage of both plagiarism and unique content.	Percentage of both plagiarism and unique content.

Results: The study found significant variation in OSI($p < 0.001$) among the four software, with Grammarly having the highest mean rank(3.56) and Small SEO Tools having the lowest(1.67). Pairwise analyses revealed significant differences($p < 0.001$) between all pairs except for Small SEO Tools-DupliChecker. Number of words showed a significant correlation with OSI for iThenticate($p < 0.05$) but not for the other three. FRES had a positive correlation, and GFI had a negative correlation with OSI by DupliChecker. FKGL negatively correlated with OSI by Small SEO Tools and DupliChecker.

Conclusion: Grammarly is unexpectedly most effective in detecting plagiarism in AI-generated articles compared to the other tools. This could be due to different softwares using diverse data sources. This highlights the potential for lower-cost plagiarism detection tools to be utilized by researchers.

1. Introduction

In the present competitive academic era, the need to prove one's academic prowess has become rampant. One way to achieve this ambition is by publishing quick and many research publications. During the COVID-19 pandemic, there was a significant surge in the number of articles being published, primarily attributed to the increased flexibility in journal acceptance policies. As a result, numerous articles were published, not only those pertaining to COVID-19 [1,2] but also on a wide range of other topics [3,4]. Due to this rush in the publishing trend, many commit plagiarism knowingly or unknowingly. The Encyclopedia Britannica defines plagiarism as "the act of taking the writings of another person and passing them off as one's own" [5]. It can be the theft of an idea, text, or data [6–8]. The repercussion of plagiarism ranges from revision of the article, retraction of the article, suspension or publication ban of the author for a specific period [9,10]. This can lead to loss of reputation, faith among peers, or job opportunities. As a result, it is essential to ensure that any work submitted is free from plagiarism [11].

Before the digital age, the responsibility of detecting plagiarism rested on the shoulders of the reviewers and editors based solely on their ability and/or experience [11]. Nowadays, many paid or free tools are available online to detect plagiarism. With numerous options available, it is challenging to select software that provides accurate results, has an extensive vocabulary, and identifies even the slightest hint of plagiarism. One of the critical features of any plagiarism detection software is the ability to identify and highlight plagiarized content. Advanced software utilizes semantic analysis and artificial intelligence to identify paraphrased content and detect subtle forms of plagiarism.

Software with a limited vocabulary may not be able to detect plagiarism accurately. Therefore, the best software should have an extensive vocabulary and a vast database of sources to compare against the submitted work. Most paid software offers a more extensive database of sources and a more extensive vocabulary, enabling it to provide more accurate results. In contrast, free software may have a limited vocabulary, making it challenging to detect plagiarism accurately. While some free software may use the same algorithms as paid software, they may not have access to the same resources or databases. This limitation can result in missed instances of plagiarism or false positives.

The iThenticate is a paid platform, the most commonly used software by academic publishers and researchers [11]. It compares content against 97% of the top 10,000 cited journals [12]. The other paid software, Grammarly, compares the text to over 16 billion web pages and academic papers stored in ProQuest's databases [13]. It also provides information about errors in grammar, spelling, punctuation, etc., along with plagiarism detection. The Small SEO tools and DupliChecker are available in both free and paid versions and claim to compare the work against billions and millions of web pages on the internet, respectively [14,15]. For more details on the included software, see Table 1.

It is not possible to check plagiarism from already published research articles, books, or any research writings, doing so would inevitably yield complete plagiarism. Therefore, articles were created by Chat Generative Pre-Trained Transformer (ChatGPT) [16], an Artificial intelligence (AI) software. AI algorithms can analyze vast amounts of data and extract relevant information, which can be used to create an outline or draft of the article. AI for research article writing has recently gained prominence in the academic world, offering an innovative and efficient approach to producing high-quality scholarly works.

This is the first of its kind study in this field aiming to investigate the difference in the Overall Similarity Index (OSI) detected by iThenticate, a widely recognized software for detecting plagiarism in research articles, compared to the alternative software options, such as Grammarly, a paid platform, and free software such as Small SEO tools and DupliChecker. We have assessed the correlation of OSI with the number and the complexity of the words. The importance of this study lies in the fact that the subscription cost of iThenticate is substantially higher than that of other available options. Thus to promote inclusivity in research and and to allow equitable opportunities to researchers from diverse backgrounds, we have conducted this study to explore the comparative advantages of these low-cost or free alternatives, researchers could access the appropriate tools for plagiarism detection.

2. Materials and methods

2.1. Study design

This is a prospective cross-over study conducted in the department of pharmacology, All India Institute of Medical Sciences

(AIIMS), Jodhpur.

2.2. Inclusion and exclusion criteria

AI-generated articles from ChatGPT with less than 1000 words were included in the study. Research articles, grey literature, protocols, and any published works were all excluded from the study.

2.3. Aim and objectives

The aim of this study is to compare the performance of four plagiarism detection software tools namely, iThenticate, Grammarly, Small SEO tools, and DupliChecker in detecting plagiarism from 100 AI generated articles of less than 1000 words.

The primary objective is to compare the OSI with the four software tools that identify plagiarism. This will help to determine which software is most effective in detecting plagiarism in AI-generated articles. Plagiarism is represented as the overall similarity index which compares a document with the entire internet or all the databases available.

Secondary objective is to find the correlation between overall similarity indices that are obtained by the individual software and other variables, such as the length of the article or the complexity of the language used. This will help to assess how different factors might impact the plagiarism detection rates of the software.

2.4. Study procedure

A total of 10 general domains that affect various aspects of life were selected randomly such as animal use, cosmetics and pharmaceutical industry, cosmology, engineering, environment, evolution of sports, finance, gender roles, medical technology, and recent advances in chronic diseases. For each general domain, ten articles on varied topics under that domain were randomly decided, and ChatGPT was commanded to generate articles. The command given to ChatGPT was "Make new articles for 'topic name' under 1000 words". Due to the word limit imposed by the free software, which was capped at 1000 words, it was concluded that all articles should be created within 1000 words. The results generated in the chat window were copied and saved as Microsoft word document. All 100 articles were then run through the four software namely, iThenticate, Grammarly, Small SEO tools, and DupliChecker. The two free software, Small SEO tools and DupliChecker were chosen, as they were the top results in the Google search engine upon the command "free plagiarism detection tool". iThenticate and Grammarly, the most commonly used tools were accessed through the institution license, while free versions were used to access the Small SEO tools and DupliChecker. As per the University Grants Commision of India, plagiarism is categorized into four levels in ascending order of severity as follows: Level 0 (up to 10% similarity), Level 1 (above 10%–40% similarity), Level 2 (above 40%–60% similarity), and Level 3 (above 60%) [17]. The Overall Similarity Index (OSI) was recorded by four independent candidates for each software individually who were blinded to the study. The number of words was noted from the Microsoft word document of each article. The readability and complexity of words were assessed using Flesch Reading Ease Score (FRES), Flesch-Kincaid Grade Level (FKGL), and Gunning-Fog Index (GFI). The above three parameters were calculated using an online analysis tool named, WebFX [18], which was selected based on a previous study for readability [19].

The FRES serves as the established gauge of readability adopted by the US Department of Defense, providing a rating scale of 0–100 [20]. The greater the score, the easier the readability [16]. FRES and FKGL are widely used scoring systems to evaluate text readability, promoting extensive cross-comparison and generalizability of outcomes. GFI was chosen for its ability to offer an added layer of subtlety in assessing readability [21]. This index takes into account the complexity and unfamiliarity of words, utilizing a list of commonly used words that, despite possessing a relatively high number of syllables, are not classified as complex [19]. FKGL and GFI scores generate results as the educational grade levels of the United States. The first grade in the American education system typically pertains to students between the ages of 6 and 7, whereas the twelfth grade is generally associated with individuals aged 17 to 18 [22]. No scoring system has been widely accepted as the gold standard for evaluating the ease of readability. Hence, the utilization of numerous verified scoring mechanisms in this manner enables a fair and comprehensive assessment [19].

2.5. Statistical analysis

The sample size for the number of articles was calculated to be 100 based on the previous study that dealt with the plagiarism of the articles [23]. The data was carefully analyzed to ensure completeness and accuracy. To assess the normality of the data, a Kolmogorov-Smirnov test was conducted alongside a visual inspection of histograms, Q-Q plots, and stem and leaf plots. Appropriate descriptive statistics were then applied to summarize the data, with normally distributed data being interpreted as Mean $\pm$ Standard Deviation (SD), and non-normally distributed data being interpreted as Median (Interquartile range (IQR) or Range).

For non-normally distributed data, repeated-measures Friedman's analysis of variance by ranks was utilized to establish significant differences in the Overall Similarity Index (OSI) for the 100 articles across all the four software. As Friedman's test is an omnibus test, pairwise comparisons were carried out in case of a significant difference. To evaluate the multiplicity adjustment, Wilcoxon rank sum test with continuity correction, utilizing Bonferroni's method was used. Additionally, to assess the significant difference in the distribution of OSI across the ten general domains, the Kruskal Wallis test was performed.

To establish the correlation between the OSI generated by the software with the three scoring systems for readability and complexity of words (FRES, FKGL and GFI) and number of words, the Spearman's rank correlation coefficient (ρ) was used. To characterize the degree of correlation, the absolute values of rho were used as a reference. Correlations falling within the ranges of

Table 2
Topic of the domains and articles included in the study.

Sl. No	Domain	Article
1.	Cosmetics and Pharma industry	Social media influence
		Animals testing for cosmetic product
		The rise of natural and organic ingredient
		Cosmetic industry
		Product diversity
		Innovation
		Globalization
		Suatainability
		Health
		Technology and innovation
2.	Engineering	Life saving technology
		Nanotechnology
		Advanced communication
		Artificial intelligence
		Engineering in disaster response
		Engineering in water quality improvement
		Sustainable infrastructure
		Exploration of space
		Safety and security
		Economic growth
3.	Environment	Climate change
		E Waste
		Environment education
		Plastic waste
		Renewable energy
		Biodiversity
		Pollution
		Circular economy
		Sustainable tourism
		Reduce, Reuse, and Recycle: A Guide to Sustainable Living
4.	Gender roles	Role and constructs
		Intersectionality
		Education
		Inequality
		Gender pay gap
		Violence
		Gender dressing
		Socialization
		Gender and mental health
		Stereotypes
5.	Animal Use	Agriculture
		Research
		Clothing
		Entertainment
		Transport
		Therapy
		Hunting
		Tourism
		Security
		Conservation
6.	Cosmology	Big bang
		Dark matter
		Beyond earth
		Role in human culture
		Fate of universe
		Tech advancement in cosmology
		Black hole
		Galaxy formation
		Extra terrestrial life
		Cosmic waves
7.	Evolution of sports	Athletic performance
		Training technique
		Sports equipment
		Sports medicine
		Sports psychology
		Sports nutrition
		Sports league
		Tourism

Table 2 (*continued*)

Sl. No	Domain	Article
8.	Finance	Women sports
		From niche hobby to mainstream
		Olympic games
		SIP Vs Mutual funds
		Mutual funds Vs Equity
		Pension pros and Cons
		Inflation
		Forex
		Credit score
		Real estate
		Banking
		Insurance
		Shopping
9.	Medical Technology	Gene editing potential
		AI in healthcare transformation
		Robotic surgery improvements
		Nanotechnology in cancer treatment
		Wireless medical devices
		3D printing
		Wearable tech in healthcare
		Telemedicine is transforming healthcare
		VR in pain management
		Blockchain in healthcare
10.	Recent advances in chronic diseases	Diabetes
		Hypertension
		Heart failure
		Ischemic heart disease
		Osteoarthritis
		Asthma
		HIV
		Neurodegenerative
		Obesity
		CKD

0–0.19, 0.2–0.39, 0.40–0.59, 0.6–0.79, and 0.8–1 were considered to be very weak, weak, moderate, strong, and very strong, respectively [24]. SPSS version 25 and R version 4.2.1 were employed for statistical analysis and visualization. The p-value of less than 0.05 was considered statistically significant.

3. Results

A total of 100 articles generated by an AI software, ChatGPT were included in the study (See Table 2). The study flowchart has been illustrated in Supplementary Fig. S1. The overall similarity index data for each software was found to be non-normally distributed (See Supplementary Fig. S2). The number of words and complexity of words indices were found to be normally distributed. The Median (IQR) for OSI generated by Grammarly was 8.50 (4–13) being the highest, while iThenticate, DupliChecker and Small SEO tools had Median (IQR) of 4 (0–11.75), 0 (0–4), and 0 (0–2.75), respectively (See Fig. 1). The OSI values produced by iThenticate, Grammarly,
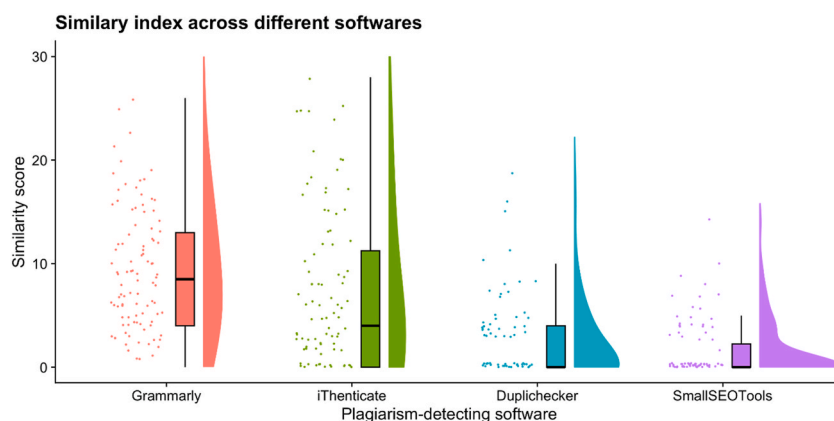

**Fig. 1.** Overall Similarity Index of articles across the 4 plagiarism detection software.

and Small SEO tools exhibited significant variations among the 10 general domains (p < 0.05). In contrast, DupliChecker did not yield any noticeable difference among the groups (p > 0.05) (See Fig. 2A, B, 2C, 2D and Supplementary Table S1).

There was a significant difference in the OSI reported by the software tools (p < 0.001), with Grammarly reporting the highest mean rank of 3.56 and Small SEO tools reporting the lowest mean rank of 1.67. The other software tools fell in between, with iThenticate and DupliChecker reporting mean rank of 2.87 and 1.91, respectively (See Supplementary Fig. S3). Pairwise analyses of the four-software found that all the five pairs were significantly different (p < 0.001) except Small SEO Tools - DupliChecker pair (p = 1.00) (See Fig. 3).

There were significant differences between the number of words, scores for readability and complexity of words (FRES, FKGL, and GFI) among the domains of the articles (p values < 0.001). The average number of words, FRES, FKGL and GFI among the domains were 512 ± 88.79, 33.379 ± 13.319, 13.597 ± 2.167, and 17.283 ± 2.782 (Mean ± SD), respectively (See Supplementary Table S2). For iThenticate, the correlation between the number of words and the OSI ($\rho = 0.12$, p value = 0.048) indicated a very weak positive significant relationship between the two variables. While there was no significant correlation between OSI and the three scores namely, FRES, FKGL, and GFI. In case of Grammarly, there was no statistically significant correlation of OSI with number of words, FRES, FKGL, and GFI. The correlation between OSI and FKGL was weak, negative, and significant ($\rho = -0.218$, p = 0.030) in Small SEO tools. The OSI by Small SEO tools showed no statistically significant correlation with number of words, FRES and GFI. There was a significant weak positive correlation between FRES and OSI by DupliChecker ($\rho = 0.211$, p = 0.035). FKGL and GFI had a significant weak negative correlation with the OSI by DupliChecker ($\rho = -0.285$, p = 0.004) and ($\rho = -0.236$, p = 0.008). There was non-significant correlation between the number of words and the ability of DupliChecker to detect plagiarism (See Supplementary Table S3).

4. Discussion

In the online education and research era, plagiarism has become a serious concern for educators, researchers, and students alike. To combat this issue, many plagiarism software tools have emerged in the market, each claiming to be the best. However, finding the right software to meet one's specific needs can be daunting. This is a first of its kind study that evaluated the performance or efficiency of
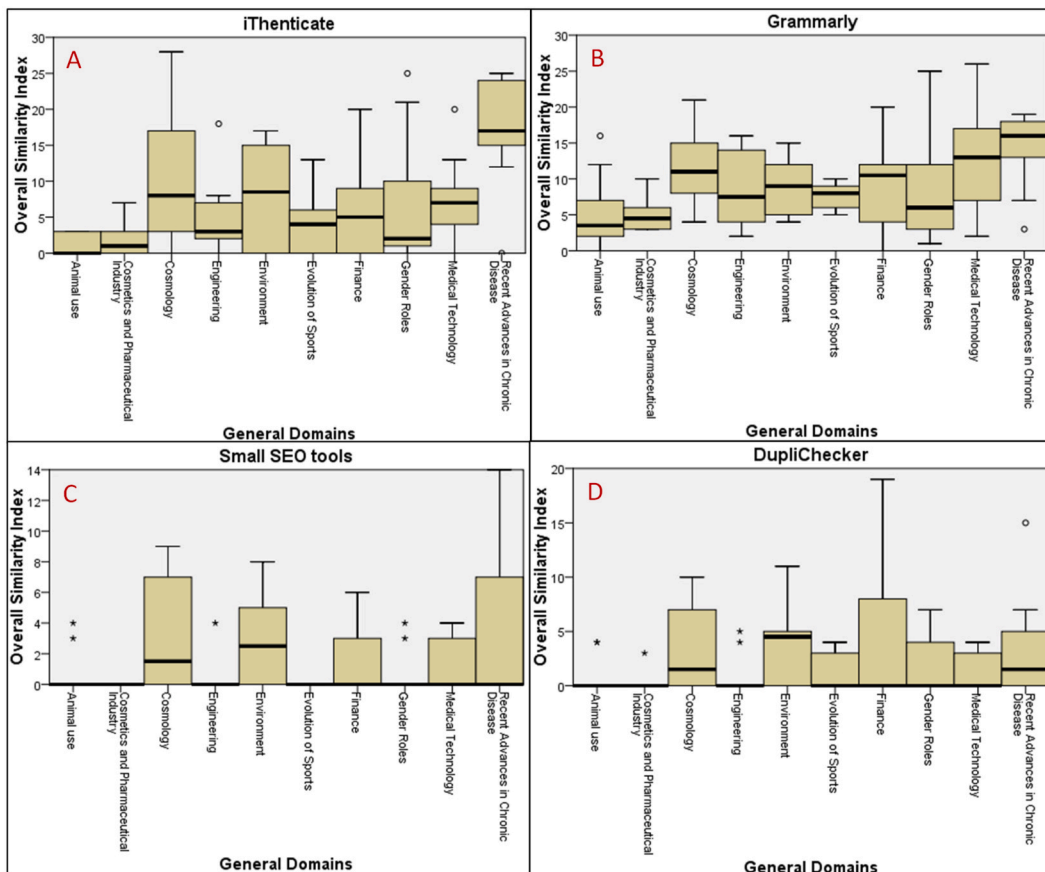
A). iThenticate, B). Grammarly, C). Small SEO tools and D). DupliChecker

Fig. 2. Distribution of Overall Similarity Index among different general domains.

Pairwise Comparisons

Overall Similarity Index (Grammarly)
3.56

Overall Similarity Index (iThenticate)
2.87

Overall Similarity Index (DupliChecker)
1.91

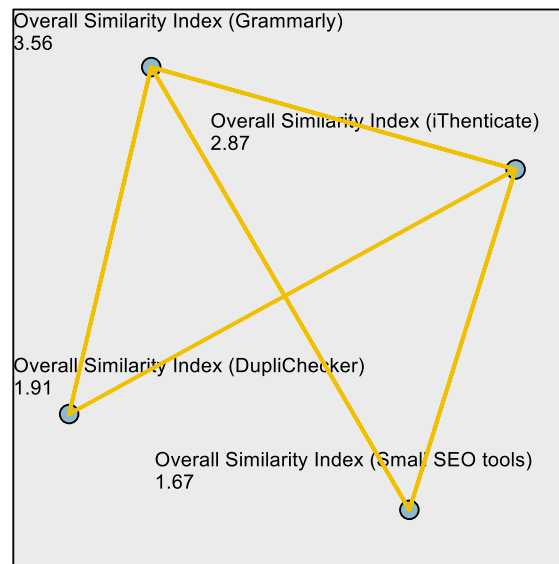Overall Similarity Index (Small SEO tools)
1.67

Fig. 3. Mean rank and pairwise comparison of the Overall Similarity Index across the four software.

four plagiarism-checking software, iThenticate, Grammarly, Small SEO tools, and DupliChecker, using articles generated by ChatGPT.

Upon analysis, it was found that Grammarly had the highest mean rank of OSI, followed by iThenticate. The two free software had a nearly identical and relatively low mean rank compared to the two paid software. From the pairwise comparison of the four software, it was found that all the pairs were significantly different, except for the Small SEO tools - DupliChecker pair, which showed no difference statistically. The above findings suggest that Grammarly is better at detecting plagiarism in AI-generated texts compared to the other three software.

It is important to remember that the overall similarity index does not equate to the plagiarism index. Although there isn't a set benchmark for what constitutes a good or bad similarity score, it's crucial to recognize that a 0% similarity score does not necessarily signify a perfect paper, and a high similarity score does not always mean that the researcher has committed plagiarism [25]. The similarity index is the percentage of similarity when a document is compared with a set of documents. In contrast, the overall similarity index is the comparison with the entire internet or all the databases available [26,27]. The main strength of the study is that the incidence of false positives was avoided by manually looking for the matching of the highlighted plagiarized contents in the reports by reviewing every article across the four software platforms. Utilizing AI for article generation played a key role in the study, as the published articles and books will show an overall similarity index of nearly 100% which could render the study infeasible.

The number of words may have some influence on the overall similarity index of iThenticate, but the relationship was not strong and may depend on other factors. There was no relationship between the overall similarity index of Grammarly and the number of words or the three readability and complexity scores (FRES, FKGL, and GFI). The Small SEO tools had a complex correlation by not being correlated with the number of words and FRES, while negatively correlated with FKGL and GFI. DupliChecker was positively correlated with FRES, while the correlation was negative with FKGL and GFI. This indicates that plagiarism in articles with higher FRES scores is more likely detected by DupliChecker, while suggesting that articles with lower FKGL and GFI scores are more likely to be detected as plagiarized by DupliChecker.

The results suggest that iThenticate and Grammarly may be the most suitable options for detecting plagiarism based on the articles generated by ChatGPT. While Grammarly and iThenticate showed a comparatively more uniform distribution, Small SEO tools and DupliChecker had highly skewed data, with most of the observations at zero. Both the median and the maximum values for Grammarly and iThenticate were higher than the other two.

While free plagiarism detection software may seem like an attractive option, it may not provide the same level of accuracy, advanced detection techniques, and extensive vocabulary as paid software. Users must consider their specific needs and requirements when choosing plagiarism detection software. Investing in paid software may provide more accurate results and additional features, making it a worthwhile investment in the long run. Cost is also an important consideration. While there are many free plagiarism software options available, they may not always offer the same level of accuracy and features as paid versions. However, the cost of the software should be reasonable and within the budget of the target audience.

To combat plagiarism effectively, online plagiarism checker software can be used before submitting the thesis and manuscripts, and by carefully citing the source. Researchers can enhance the sensitivity of plagiarism detection by utilizing both iThenticate and Grammarly tools together. iThenticate effectively identifies similarity by its database of published literature, while Grammarly excels

at detecting similarities among online resources. Therefore, the combined use of these tools can significantly improve the effectiveness of plagiarism detection. Suggestions to avoid plagiarism in articles [28,29] have been provided in Table 3.

4.1. Limitations of the study

The limitations of the study include the inability to utilize articles created by humans for more accurate plagiarism detection, as this study lacked support from journals or plagiarism detection software with access to unpublished manuscripts. Hence, the research can be extended towards evaluating the effectiveness of software-based plagiarism detection techniques in scrutinizing articles written by humans along with expanding the study to a wide range of plagiarism detection tools. Though research articles are not completely AI-generated, there is a clear trend towards greater use of AI tools for medical writing. Additionally, despite instructions to create articles with 1000 words, ChatGPT provided articles with fewer words for some articles, resulting in a substantial difference in the number of words among the 100 articles.

5. Conclusion

Grammarly is the most effective in detecting plagiarism in AI-generated articles, followed by iThenticate, Small SEO tools and DupliChecker. Given the cost variation and additional features present in the different options, researchers can reevaluate their choice of similarity detection software based on this research. By establishing the efficacy of lower-cost alternatives, researchers could access the appropriate tools for detecting plagiarism with ease. When feasible, Grammarly and iThenticate can both complement each other due to their vastly different similarity detection methods.

Disclaimer

ChatGPT was not used for data analysis and result interpretation.

Ethics approval

Not applicable.

Consent to participate

Not applicable.

Consent for publication

All the authors contributed in the final approval of the manuscript and hereby approve for the publication of the work.

Declarations

Author contribution statement

Ranjit Sah: Analyzed and interpreted the data; Contributed analysis tools or data.
Abhishek Anil: Conceived and designed the experiment; Performed the experiment; Analyzed and interpreted the data; Contributed analysis tools or data; Wrote the paper.
Aswini Saravanan: Conceived and designed the experiment; Performed the experiment; Analyzed and interpreted the data; Contributed analysis tools or data; Wrote the paper.
Surjit Singh: Conceived and designed the experiment; Analyzed and interpreted the data; Contributed analysis tools or data; Wrote the paper.
Muhammad Aaqib Shamim: Conceived and designed the experiment; Analyzed and interpreted the data; Contributed analysis tools or data; Wrote the paper.
Krishna Tiwari: Performed the experiment; Analyzed and interpreted the data; Contributed analysis tools or data; Wrote the paper.
Hina Lal: Performed the experiment; Analyzed and interpreted the data; Contributed analysis tools or data; Wrote the paper.
Shanmugapriya Seshatri: Performed the experiment; Analyzed and interpreted the data; Contributed analysis tools or data.
Simi Bridjit Gomaz: Performed the experiment; Analyzed and interpreted the data; Contributed analysis tools or data; Wrote the paper.
Thoyyib P Karat: Conceived and designed the experiment; Analyzed and interpreted the data; Contributed analysis tools or data; Wrote the paper.
Pradeep Dwivedi: Conceived and designed the experiment; Analyzed and interpreted the data; Contributed analysis tools or data; Wrote the paper.
Shoban Babu Varthya: Conceived and designed the experiment; Analyzed and interpreted the data; Contributed analysis tools or data; Wrote the paper.

Table 3
Suggestions to avoid plagiarism in articles.

- To prevent plagiarism, it is imperative to possess a clear understanding of its definition.
- Online plagiarism checker tools are readily available to check one's work for potential instances of plagiarism.
- It is important to utilize one's individual words while crafting articles.
- Despite the fact that one may be referencing the notions of others, it is crucial to rephrase them using one's own diction.
- It is essential to acknowledge and attribute the ideas and words of others when incorporating them into one's work, which can be achieved by citing the source.
- Paraphrasing entails the art of articulating another individual's concepts in one's own diction, serving as a pragmatic technique to sidestep plagiarism while simultaneously incorporating their ideas.
- When incorporating verbatim content from another individual's work, quotation marks should be used to indicate that the text is a direct quotation. The quote must be attributed to its rightful author and cite the source from which it was derived.
- Incorporating diverse sources into your research endeavours can prove advantageous in circumventing inadvertent plagiarism. This approach offers a better comprehension of the subject matter and diminishes the probability of duplicating someone else's ideas.
- Effective preparation can assist in evading plagiarism. It is recommended to strategize your writing in advance and guarantee that you allocate sufficient time to explore and pen your thoughts without hurrying.
- General knowledge refers to information that is widely recognized and acknowledged as factual. It is not necessary to provide a citation for general knowledge.
- It is paramount to maintain veracity in your writing. The act of plagiarism carries severe ramifications; hence it is advisable to be forthright and truthful.

Rimple Jeet Kaur: Conceived and designed the experiment; Analyzed and interpreted the data; Contributed analysis tools or data.
Prakasini Satapathy: Analyzed and interpreted the data; Contributed analysis tools or data.
Bijaya Kumar Padhi: Analyzed and interpreted the data.
Shilpa Gaidhane: Analyzed and interpreted the data; Wrote the paper.
Manoj Patil: Analyzed and interpreted the data; Wrote the paper.
Mahalaqua Nazli Khatib: Analyzed and interpreted the data; Wrote the paper.
Joshuan J. Barboza: Analyzed and interpreted the data; Contributed analysis tools or data.

Data availability statement

Data will be available on reasonable request.

Research in context

Evidence before this study

- While there have been studies on the effectiveness of plagiarism detection software, none have statistically assessed its efficacy, nor have any incorporated AI-generated articles into their research.
- Furthermore, the impact of article length and language complexity on plagiarism detection remains unexplored.

Added value of this study

- This study compares the effectiveness of paid and free plagiarism detection software tools in identifying plagiarism in AI-generated articles, marking the first of its kind.
- The results reveal that Grammarly outperforms the other tools in detecting plagiarism in terms of OSI, especially iThenticate which is the most commonly used tool and also has a subscription cost higher than the other tools.
- It also sheds light on the influence of article length and language complexity on plagiarism detection.

Implication of the available evidence

- This study offers valuable insights that can assist researchers in selecting the most suitable tools for their work.
- The significance of the study lies in its potential to revolutionize plagiarism detection, making it more accessible to researchers and paving the way for future advancements in the field.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e19194

References

[1] Antai college of economics and management, school of media and communication, Shanghai jiao tong university, minhang district, Shanghai, China, abbas J. the impact of coronavirus (SARS-COV2) epidemic on individuals mental health: the protective measures of Pakistan in managing and sustaining transmissible disease, Psychiatr. Danub. 32 (3–4) (2020 Dec 24) 472–477.

[2] J. Abbas, D. Wang, Z. Su, A. Ziapour, The role of social media in the advent of COVID-19 pandemic: crisis management, mental health challenges and implications, Risk Manag. Healthc. Pol. 14 (2021 May) 1917–1932.

[3] C.A. Schmidt, E.A. Cromwell, E. Hill, K.M. Donkers, M.F. Schipp, K.B. Johnson, et al., The prevalence of onchocerciasis in Africa and Yemen, 2000–2018: a geospatial analysis, BMC Med. 20 (1) (2022 Sep 7) 293.

[4] P.T. Iorember, B. Iormom, T.P. Jato, J. Abbas, Understanding the bearable link between ecology and health outcomes: the criticality of human capital development and energy use, Heliyon 8 (12) (2022 Dec), e12611.

[5] Plagiarism | Britannica [internet]. 2023 [accessed 2023 mar 16], Available from: https://www.britannica.com/topic/plagiarism.

[6] C.T. Ambrose, Plagiarism of ideas. Benjamin rush and charles caldwell–a student-mentor dispute, Pharos Alpha Omega Alpha-Honor Med Soc Alpha Omega Alpha 77 (1) (2014) 14–23.

[7] N. Das, M. Panjabi, Plagiarism: why is it such a big issue for medical writers? Perspect Clin Res 2 (2) (2011) 67–71.

[8] B.H. Reyes, Plagiarism in scientific publications, Rev Med Chil 137 (1) (2009 Jan) 7–9.

[9] S.S. Khadilkar, The plague of plagiarism: prevention and cure!!!, J Obstet Gynecol India 68 (6) (2018 Dec 1) 425–431.

[10] H. Sharma, S. Verma, Insight into modern-day plagiarism: the science of pseudo research, Tzu Chi Med. J. 32 (3) (2019 Dec 5) 240–244.

[11] Y.B. Roka, Plagiarism: types, causes and how to avoid this worldwide problem, Nepal J. Neurosci. 14 (3) (2017 Dec 1) 2–6.

[12] Turnitin iThenticate by. Plagiarism detection software | iThenticate [internet] [accessed 2023 Mar 26]. Available from: https://www.ithenticate.com.

[13] About grammarly – grammarly support [internet] [accessed 2023 Mar 25]. Available from: https://support.grammarly.com/hc/en-us/categories/115000018611-About-Grammarly.

[14] Plagiarism checker - 100% free plagiarism detector online [internet]. Plagiarism checker - 100% free plagiarism detector online. [accessed 2023 mar 25], Available from: https://smallseotools.com/plagiarism-checker/.

[15] Plagiarism checker | 100% free and accurate - duplichecker [internet]. Duplichecker.com. [accessed 2023 mar 25], Available from: https://www.duplichecker.com/.

[16] ChatGPT [internet] [accessed 2023 Mar 25]. Available from: https://chat.openai.com.

[17] [cited 2023 Jul 10], Available from: https://www.ugc.gov.in/pdfnews/7771545_academic- integrity-Regulation2018.pdf.

[18] Readability test [internet]. WebFX. [accessed 2023 Mar 25], Available from: https://www.webfx.com/tools/read-able/.

[19] N.E. Wrigley Kelly, K.E. Murray, C. McCarthy, D.B. O'Shea, An objective analysis of quality and readability of online information on COVID-19, Health Technol. 11 (5) (2021 Sep) 1093–1099.

[20] L. Si, J. Callan, A statistical model for scientific readability [accessed 2023 Mar 18]. pp. 574–6. Available from:, in: Proceedings of the Tenth International Conference on Information and Knowledge Management [Internet], ACM, Atlanta Georgia USA, 2001 https://dl.acm.org/doi/10.1145/502585.502695.

[21] J. Kincaid, R. Fishburne, R. Rogers, B. Chissom, Derivation of new readability formulas (automated readability index, Fog count and Flesch reading ease formula) for navy enlisted personnel. Inst simul train [internet] [accessed 2023 Mar 20] 1975 Jan 1; Available from: https://stars.library.ucf.edu/istlibrary/56.

[22] Structure of U.S. Education [internet] [accessed 2023 Mar 18]. Available from: https://www2.ed.gov/about/offices/list/ous/international/usnei/us/edlite-structure-us.html.

[23] A. Rohwer, E. Wager, T. Young, P. Garner, Plagiarism in research: a survey of African medical journals, BMJ Open 8 (11) (2018 Nov), e024777.

[24] 11. Correlation and regression | the BMJ [Internet]. The BMJ | the BMJ: leading general medical journal. Research. Education, Comment. 2020 [accessed 2023 Mar 19]. Available from: https://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one/11-correlation-and-regression.

[25] S.A. Meo, M. Talha, Turnitin: is it a text matching or plagiarism detection tool? Saudi J. Anaesth. 13 (5) (2019 Apr 1) 48.

[26] E. Stamatatos, A survey of modern authorship attribution methods, J. Am. Soc. Inf. Sci. Technol. 60 (3) (2009 Mar) 538–556.

[27] Y. Guodong, Y. Yu, Z. Xi, Liu-Aijun, Evolution model of collaborative production task responding to unplanned perturbations, J. Intell. Fuzzy Syst. 30 (1) (2015 Oct 6) 383–396.

[28] Pm Kumar, Ns Priya, S. Musalaiah, M. Nagasree, Knowing and avoiding plagiarism during scientific writing, Ann. Med. Health Sci. Res. 4 (9) (2014) 193.

[29] Rg Steen, Writing for publication in a medical journal, Indian J Endocrinol Metab 16 (6) (2012) 899.