



Original article

Detecting generative artificial intelligence in scientific articles: Evasion techniques and implications for scientific integrity

Guillaume-Anthony Odri^{a,b,*}, Diane Ji Yun Yoon^b^a Service de chirurgie orthopédique et traumatologique, centre hospitalier universitaire Lariboisière, 2, rue Ambroise-Paré, 75010 Paris, France^b Inserm U1132 BIOSCAR, université Paris-Cité, 75010 Paris, France

ARTICLE INFO

Article history:

Received 15 July 2023

Accepted 19 September 2023

Keywords:

Generative artificial intelligence

Academic writing

Scientific fraud

ABSTRACT

Background: Artificial intelligence (AI) tools, although beneficial for data collection and analysis, can also facilitate scientific fraud. AI detectors can help resolve this problem, but their effectiveness depends on their ability to track AI progress. In addition, many methods of evading AI detection exist and their constantly evolving sophistication can make the task more difficult. Thus, from an AI-generated text, we wanted to: (1) evaluate the AI detection sites on a text generated entirely by the AI, (2) test the methods described for evading AI detection, and (3) evaluate the effectiveness of these methods to evade AI detection on the sites tested previously.

Hypothesis: Not all AI detection tools are equally effective in detecting AI-generated text and some techniques used to evade AI detection can make an AI-produced text almost undetectable.

Materials and methods: We created a text with ChatGPT-4 (Chat Generative Pre-trained Transformer) and submitted it to 11 AI detection web tools (Originality, ZeroGPT, Writer, Copyleaks, Crossplag, GPTZero, Sapling, Content at scale, Corrector, Writefull et Quill), before and after applying strategies to minimise AI detection. The strategies used to minimize AI detection were the improvement of command messages in ChatPGT, the introduction of minor grammatical errors such as comma deletion, paraphrasing, and the substitution of Latin letters with similar Cyrillic letters (a and o) which is also a method used elsewhere to evade the detection of plagiarism. We have also tested the effectiveness of these tools in correctly identifying a scientific text written by a human in 1960.

Results: From the initial text generated by the AI, 7 of the 11 detectors concluded that the text was mainly written by humans. Subsequently, the introduction of simple modifications, such as the removal of commas or paraphrasing can effectively reduce AI detection and make the text appear human for all detectors. In addition, replacing certain Latin letters with Cyrillic letters can make an AI text completely undetectable. Finally, we observe that in a paradoxical way, certain sites detect a significant proportion of AI in a text written by a human in 1960.

Discussion: AI detectors have low efficiency, and simple modifications can allow even the most robust detectors to be easily bypassed. The rapid development of generative AI raises questions about the future of scientific writing but also about the detection of scientific fraud, such as data fabrication.

Level of evidence: III Control case study.

© 2023 Elsevier Masson SAS. All rights reserved.

1. Introduction

The expression “publish or perish” represents an omnipresent reality in the academic and scientific world, reflecting the intense pressure exerted on researchers to regularly produce and publish

research results [1]. This pressure can unfortunately lead to deviant practices, going as far as scientific fraud, which manifests itself in various ways such as data fabrication and plagiarism [2]. In this context, the advent of artificial intelligence (AI), particularly tools such as GPT-4 (Generative Pre-trained Transformer-4), and the conversational agent ChatGPT, presents both opportunities and challenges [3].

On the one hand, these AI tools can be valuable for scientific research, aiding data collection and analysis, as well as writing of draft articles, or prompting new avenues for research. On the other hand, they can be used in an abusive manner, and thus facil-

* Corresponding author at: Service de chirurgie orthopédique et traumatologique, centre hospitalier universitaire Lariboisière, 2, rue Ambroise-Paré, 75010 Paris, France.

E-mail address: guillaume.odri@aphp.fr (G.-A. Odri).

"Writing scientific articles is a complex task, requiring expertise in a specific field and the ability to present information clearly and concisely. Sometimes, artificial intelligence tools such as Chat GPT can be used to assist in this task. However, it is crucial to ensure that the text generated is not easily identifiable as having been produced by an AI. Several AI detection software packages, such as Originality.ai, are very good at detecting AI-generated text, which makes this challenge all the more important.

It should be stressed that this method should not be used to create a complete article just because it has been written by a human. Scientific journal editors and ethics committees have strict rules on this. However, if one is using Chat GPT to complete one's own editorial work, it can be useful to understand how to minimize detection by AI.

Three main strategies can help produce content generated by Chat GPT that more closely resembles text written by a human: increasing "burstiness", "perplexity" and "temperature".

1. ****Increase "Burstiness"**: This method relies on introducing variability into the generated text, making the writing style more similar to that of humans.**
2. ****Increase "Perplexity"**: Perplexity is a measure of a linguistic model's ability to predict a sequence of words. A higher level of perplexity indicates that the model has more difficulty predicting the next word, which can lead to a more diverse text.**
3. ****Increase "Temperature"**: This strategy uses a hyperparameter to control the level of creativity and diversity of the text generated by the model. A higher temperature leads to more diverse and creative text.**

It is essential to understand that even with these adjustments, this approach is not infallible. AI detection software continues to improve and can still detect AI-generated text. In an originality test carried out with Originality.ai, a text generated by GPT Chat, even after being adjusted to minimize AI detection, scored 95% originality and was judged 5% human-origin. Despite the high score, the style of the text was such that it was not taken seriously.

Therefore, it's advisable not to simply copy and paste Chat GPT's work, whether you're a student or writing content for a website. Instead, extract value from the AI-generated text to enrich your own writing.

It's also crucial to check the originality of your work. Tools like Originality.ai (which is an affiliate partner) are very accurate at detecting AI-generated text. It has been proven that Quillbot-reformulated content always fails the originality test on Originality.ai, even though it may pass on other originality-checking platforms.

In summary, when used ethically and appropriately, Chat GPT can be a valuable tool to complement the writing of a scientific article. However, it is important to understand how to minimize AI detection, to verify the originality of the work, and to always be transparent about the use of these tools. It's essential to avoid any accusations of plagiarism and to guarantee the authenticity and credibility of your work. In addition, it should be noted that some content written by humans can be falsely identified as being of AI origin, hence the importance of checking your work with an AI-generated text detector."

Fig. 1. Initial text (Text 1) produced by GPT-4 through the ChatGPT interface.

itate scientific fraud. Faced with these challenges, AI detectors are emerging as essential tools to prevent and detect fraud [4,5]. AI detectors work by analyzing specific characteristics of the text such as perplexity, which measures the unpredictability of a text, and "burstiness", which assesses the variation in the structure and length of sentences. These tools use natural language processing models and automatic learning algorithms to identify typical pattern in texts generated by AI, such as word frequency, syntactic and semantic analysis, and comparing the text with common examples of texts generated by AI [6]. However, their effectiveness is intrinsically linked to their ability to keep up with the pace of AI progress and to the human inventiveness to escape it. As AI models become more sophisticated and produce text even closer to that of a human, AI detection tools should evolve in parallel.

In order to provide new data on this subject, a text was generated with GPT-4 via the ChatGPT4 interface and was submitted to several AI detection tools. Then different modifications were carried out on the text with the aim of reducing the rate of AI detected, and these modified texts were resubmitted to AI detectors. This study was intended to: (1) evaluate the effectiveness of AI detection sites on a text generated entirely by AI, (2) test the methods described to evade AI detection and (3) assess effectiveness of these methods to evade AI detection on the previously tested sites. Our hypothesis was that not all AI detection sites are equally effective in detecting AI-generated text and some techniques to evade AI detection could render a text produced by AI almost undetectable.

2. Materials and methods

2.1. Material

To generate the initial text written by GPT-4, we carried out a Google search in the first week of July 2023; "how to beat AI detection?" with the "last month" filter and excluding sponsored sites and videos. We copied the text of 4 sites of interest into a Word file [7–10]. We then asked ChatGPT [11], using GTP-4, to create a summary of these 4 sites in a single text, in order to have a text written by an AI describing the method of minimising AI detection. We reread and corrected the text. All work was carried out on English texts. The initial text, called Text 1, is in Fig. 1 and was generated in 12 minutes.

2.2. Methods

We then took this text, and we passed it through 11 AI detectors present on internet sites: ZeroGPT, Originality, Writer AI content Detector, Copyleaks, Crossplag, GPTZero, Sapling, Contentatscale, Corrector, Writefull and Quill. Some of these detectors are completely free, while others require payment (Originality, the paid option was therefore used), others are free but only for a limited number of characters, others offer paid options for more text and have advanced features, but it is not clearly established that the paid version is associated with a more powerful AI detector, and thus the free options were used (ZeroGPT, Copyleaks, Sapling). The following sites were used to detect AI during the first week of July 2023:

- ZeroGPT: <https://www.zerogpt.com/>, free, whose site indicates that it is the most advanced and reliable Chat GPT, GPT-4 and AI content detector;
- Originality: <https://originality.ai/>, paid, indicated to be the most precise AI and plagiarism detector for publishers of academic content, and to be superior than CPTZero, CopyLeaks end Writer;
- Writer AI content detector: <https://writer.com/ai-content-detector/>, free for a maximum of 1500 characters;
- CopyLeaks: <https://copyleaks.com/ai-content-detector>, free;
- Crossplag: <https://crossplag.com/ai-content-detector/>, free;
- GPTZero: <https://gptzero.me/>, free;
- Sapling: <https://sapling.ai/ai-content-detector>, free for up to 2000 characters;
- Content at scale: <https://contentatscale.ai/ai-content-detector/>, free;
- Corrector: <https://corrector.app/ai-content-detector/>, free;
- Writefull: <https://x.writefull.com/gpt-detector>, free;
- Quill: <https://aiwritingcheck.org/>, free for up to 400 words.

Subsequently, different techniques to try and reduce AI detection were applied, in a progressive manner (Fig. 2), and the modified text was then passed through the AI detectors. The following techniques used were:

- written recommendations in the ChatGPT text (increase in "perplexity", "burstiness" and "temperature") = text 2;
- introduction of minor grammatical errors: removal of a few commas and gerunds (-ing endings) = text 3;

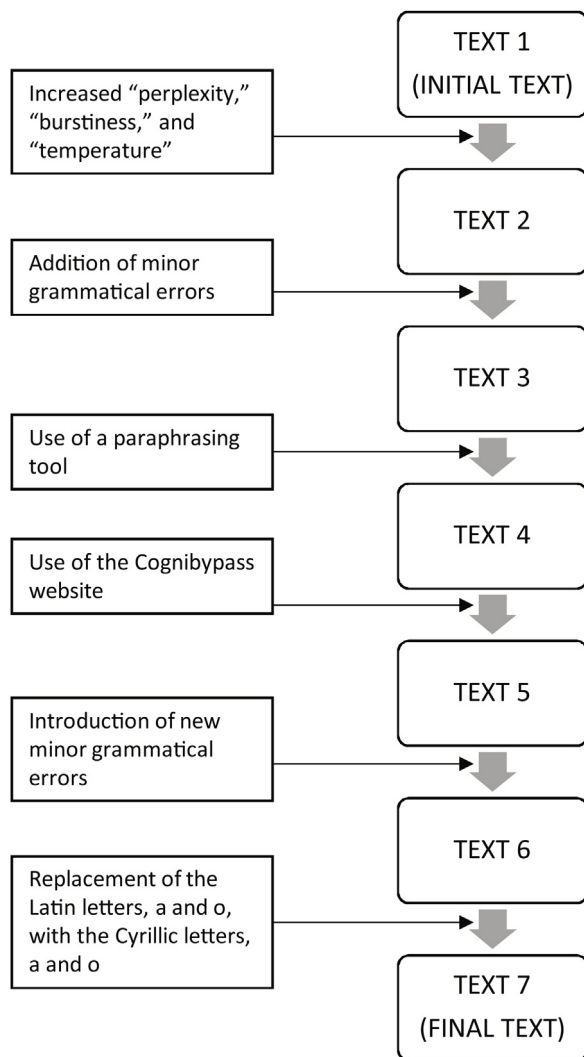


Fig. 2. Flowchart detailing text modifications ($n=6$) from the text in Fig. 1 leading to the text in Fig. 3.

- the use of a paraphrasing tool (Quillbot, <https://quillbot.com/>) = text 4;
- the use of the Cognibypass site (<https://cognibypass.com/>) which transforms the text by AI, using a complex algorithm, and which promises to beat Originality and Turnitin (<https://www.turnitin.com/>); these 2 latter detectors being considered as the best = text 5;
- introduction of new minor grammatical errors = text 6;
- replacing the Latin letters "a" and "o" with the Cyrillic letters "а" and "о" = text 7 or final text (Fig. 3). This method evades plagiarism detection. These letters look very similar, but the words in which they are inserted are not always recognized by the AI detection sites. New techniques such as these are commonly described and can be easily found on internet sites or in dedicated forums.

Completion of all these modifications took 13 minutes. We therefore dedicated a total of 25 minutes to generate the text and apply the modifications with the aim of minimizing AI detection. The final text is presented in Fig. 2. To assess the ability of these detectors to evaluate a text written by a human, we used a free-access text written by Sir John Charnley in 1960 [12]. The text in PDF format was downloaded and then converted into Word format. We eliminated all page mismatches caused by the PDF to Word conver-

sion of this text, which could influence AI detection, such as broken sentences and page breaks. The text was then passed through the 11 AI detectors.

2.3. Evaluation methods

AI detectors give different types of results: either a prediction of human origin or AI origin of the text (ZeroGPT, Copyleaks, Crossplag, GPTZero, Content at Scale, Quill), the probability (in the form of a percentage) of whether the text taken in its entirety is of human or AI origin (Originality, Copyleaks, Writefull), the proportion (in the form of a percentage) of the text of human or AI origin (ZeroGPT, Writer, Crossplag, Sapling, Content at Scale, Corrector). Some detectors highlight areas of text identified as being of AI origin (GPTZero, ZeroGPT).

2.4. Statistical methods

The results are given in percentages or proportions. No statistical validation of the results was applied.

3. Results

3.1. Initial text written by GPT4 (Text 1)

Seven detectors concluded that the text was written by a human: GPTzero, Writer (100%), Quill, Writefull (1% AI), Corrector (0.02% fake), Copyleaks (59.9% probability), and Content at scale (84% human) (Table 1). ZeroGPT predicted that the text was human but that it may have parts written by AI (25.8% by AI) (Table 1). Three detectors, Sapling (98.2% AI), Originality and Crossplag (Table 1), responded that the text had a strong probability of being written by AI, or had a high proportion of original AI text.

3.2. Text modified by ChatGPT according to the recommendations described in the text produced by ChatGPT4 (Text 2)

Nine detectors concluded that the text was written by humans, including Sapling which only found 8.6% by AI and ZeroGPT found 15.54% (Table 1). Among the previous 7, Content at scale increased to 99% human, Copyleaks remained stable (58.5% probability). For Originality and Crossplag, the text remained 100% AI (Table 1).

3.3. Text 2 + introduction of some minor grammatical errors (Text 3)

Ten detectors concluded that the text was written by a human, including Crossplag which went from 100% AI to 100% human just by this process (Table 1). For ZeroGPT, there was a decrease to 8.54% AI. For Sapling, the AI rate remained stable at 8.4%, and for Copyleaks the probability of human text was 61.2% (Table 1). Only Origin considered the text written as being written by AI with a probability of 95% (Table 1).

3.4. Text 3 + paraphrasing in Quillbot (Text 4)

The same 10 detectors concluded that the text was written by humans, with a drop to 0% AI for ZeroGPT (Table 1). For Copyleaks, the probability of the text being written by a human was 62.3%. For Originality, there was a 100% probability that the text was written by AI. For Sapling, the AI rate rose to 26.7%. For Content at scale, the text was human but at 94% (Table 1).

"The procedure of carrying out clinical study demands a high degree of knowledge in a particular domain plus the ability to efficiently connect details in a clear as well as succinct way. Expert system devices like ChatGPT can give assistance in this task. However it is important to assure that the produced message does not display obvious features that would easily determine it as being authored by an expert system. AI discovery software applications such as Originality.ai have actually shown remarkable efficiency in finding AI-generated message thus enhancing the intricacy of this concern.

It is important to recognize that this method ought to not be used for the function of creating a whole write-up with the misleading case that it was authored by a person. Rigorous standards concerning this issue are applied by clinical journal editors plus values boards. Nonetheless if ChatGPT is used as an extra device for one's very own typesetting job it can be helpful to comprehend techniques for reducing AI discovery.

There are 3 key techniques that can improve the similarity of GPT Chat-generated material to human-written message: boost "burstiness," "perplexity" plus "temperature level."

One method to improve the "burstiness" of message is by integration of irregularities right into the created material consequently providing the make-up design extra comparable to human-like expression.

Secondly boost the metric of "perplexity": perplexity works as an evaluative statistics for determining the anticipating ability of a language design in regard to producing a systematic series of words. An enhanced degree of perplexity shows that the version experiences higher difficulties in precisely anticipating succeeding words consequently leading to an extra different plus varied textual outcome.

One last technique is to improve the "temperature" : to change a hyperparameter that manages the degree of resourcefulness together with range in the message generated by the design. A raised temperature level brings about a higher level of variety coupled with creative thinking in the textual result.

It is crucial to understand that regardless of these alterations this technique is not stubborn. The recurring growth of AI discovery software application makes it possible for constant optimization enabling the discovery of message produced by AI systems. Therefore it is suggested to avoid entirely relying upon the technique of replicating as well as putting ChatGPT's outcome, no matter whether one is a pupil or participated in the production of on the internet material. Using the textual outcome produced by the AI to enhance one's very own writing is an extra sensible technique.

It is critical to confirm the credibility of one's job. Originality.ai coupled with comparable devices show a high degree of accuracy in identifying messages that have actually been produced by expert system. The paraphrased web content generated by Quillbot regularly falls short the creativity analysis on Originality.ai regardless of possibly passing analyses on various other systems for examining creativity.

In recap when utilized with moral factors to consider as well as suitable use ChatGPT can function as an useful tool for promoting the conclusion of a clinical research study's structure. It is critical to understand the techniques for decreasing AI discovery in order to examine the credibility of scholastic job. In addition it is vital to preserve openness concerning the use of such devices. Making sure the evasion of plagiarism accusations as well as supporting the credibility and also reputation of one's job is of utmost value. Moreover it deserves keeping in mind that particular material created by human beings might be mistakenly credited to AI emphasizing the importance of confirming one's job utilizing an AI-based message discovery system."

Fig. 3. Final text (Text 7) after modifications to avoid AI detection.

Table 1

Results. The initial text was generated by GPT-4 through the ChatGPT interface. Texts 2 to 6 and the final text correspond to the initial text, which was gradually modified to make the AI undetectable; one sole modification technique was added to each text. The human text is a scientific text written by Sir John Charnley in 1960 [11].

| | Text 1 (initial) | Text 2 | Text 3 | Text 4 | Text 5 | Text 6 | Text 7 (final) | Human text |
|------------------|------------------|-------------|-------------|-------------|-------------|-------------|----------------|-------------|
| GPTZero | Human | Human | Human | Human | Human | Human | Human | Human |
| Writer | 100% human | 100% human | 100% human | 100% human | 89% human | 89% human | No analysis | 98% human |
| Quill | Human | Human | Human | Human | Human | Human | Human | Human |
| Writefull | 1% AI | 1% AI | 1% AI | 2% AI | 1% AI | 1% AI | 2% AI | 13% AI |
| Corrector | 0.02% IA | 0% AI | 0% AI | 0% AI | 0% AI | 0% AI | 0% AI | 41.18% IA |
| ZeroGPT | 25.8% AI | 15.54% IA | 8.54% AI | 0% AI | 0% AI | 0% AI | 0% AI | 20.41% IA |
| Sapling | 98.2% AI | 8.6% AI | 8.4% AI | 26.7% AI | 0.1% AI | 0.1% AI | 0% AI | 27.3% AI |
| Content at scale | 84% human | 99% human | 99% human | 94% human | 100% human | 100% human | 100% human | 88% human |
| Copyleaks | 59.9% human | 58.5% human | 61.2% human | 62.3% human | 62.6% human | 62.6% human | 96.9% human | 99.9% human |
| Crossplag | 100% AI | 100% AI | 100% human | 100% human | 100% human | 100% human | 100% human | 100% human |
| Originality | 100% AI | 100% AI | 95% AI | 100% AI | 39% AI | 9% AI | 0% AI | 0% AI |

3.5. Text 4 + transformation by Cognibypass (Text 5)

The same 10 detectors consistently found the text to be written by humans, with a result of almost 100%, except for Writer, which conferred a probability of 89% (Table 1). For Originality, the probability that the text was written by AI dropped significantly to 39% (Table 1).

3.6. Text 5 + introduction of new grammatical errors (deletion of commas and gerunds) (Text 6)

All detectors concluded that the text was written by a human, with a probability of 91% for Originality and 62.6% for Copyleaks (Table 1).

3.7. Text 6 + replacement of the Latin "a" and "o" letters with the Cyrillic "а" and "о" letters (final text, or Text 7)

Ten detectors concluded that the text was written by a human, with a probability of 100% for Originality and 96.9% for Copyleaks. Only Writer did not analyze the text in this format (Table 1).

3.8. Analysis of text of human origin [12]

Rather paradoxically, some detectors observed more AI in the text of human origin, than in the text generated by ChatGPT. Writefull found 13% AI, Corrector found 41.18% AI, ZeroGPT found 20.41%

AI and Sapling found 27.3% AI. Similarly, Content at scale indicated an 88% probability that the text was of human origin. Originality found a 0% probability that the text was written by AI, as did Crossplag, Copyleaks and Writer. For GPTZero and Quill, the text was of human origin (Table 1).

4. Discussion

Our study aimed to evaluate the effectiveness of text modification techniques to avoid AI detection, using a text generated entirely by GPT-4 via the ChatGPT interface. We observed that most detectors are not effective in detecting a text generated by a generative AI, and even a human text could be identified as being generated by AI. This has also been recently studied in an extensive manner by Weber-Wulff et al. [13], who observed a high rate of false positives and false negatives, and they did not recommend the use of these detectors in an academic context. Furthermore, the techniques used to modify the text generated by the AI are simple to implement and easily accessible. Although the final text is sometimes not as well worded than the initial text, being a little less readable or less precise, it nevertheless remains accurate with respect to the initial text, without any major errors.

By analyzing the different modification techniques used to evade AI detection, we discovered a range of methods that differ in their effectiveness. Some detectors are more sensitive to certain technologies. For example, Sapling changed from an AI text to a human text when ChatGPT was used to increase "burstiness,

perplexity and temperature". These techniques confer sentences of different sizes, with more variable and less precise terms, making the text closer to human oral expression. It should be noted that these modifications are difficult to implement in a scientific text, which requires succinctness and precision, without embellishments or judgment. Similarly, the introduction of minor grammatical errors, such as the removal of commas, was very effective in inducing an error from Crossplag, going from 100% AI to 100% human by this process. Simple paraphrasing tends to increase the AI rate, and during previous testing of Quillbot, paraphrasing was evaluated by 100 scientific articles of human origin. This type of paraphrasing increased the rate of AI in 100% of cases. On the other hand, the transformation of the text by Cognibypass was very effective because even Originality concluded that the text was of human origin, while it was rather resistant to other methods used. Finally, the method of substituting Latin letters with Cyrillic letters was the most effective for all detectors. Originality was thus the most resistant detector, requiring more modifications while others, such as ZeroGPT, Sapling, Copyleaks and Crossplag lost their effectiveness as soon as simple strategies to hinder detection were implemented.

It is important to note that some of these detectors found AI in the scientific text written by Charnley in 1960. Scientific or academic writing styles tend to standardize the presentation and interpretation of interpretation in a style that is closer to the generative AI style, by limiting the perplexity and the "burstiness". This explains the large proportion of text identified as coming from generative AI in scientific texts written by humans. Likewise, in the future, repeated reading of texts generated by AI could lead to humans adopting a writing style similar to that of AI. With the continued improvement of language models and generative AI such as GPT-4 or LaMDA (Language Model for Dialogue Applications), the distinction between human-generated texts and AI-generated texts will become increasingly difficult, making the detection of generative AI use almost impossible [14].

The materials and methods, the results and the conclusion of this article were written by the authors. In contrast, the abstract, the introduction and the discussion up to this paragraph were predominantly generated by GPT-4 via ChatGPT, based on the materials and methods, and results. Various strategies to reduce AI detection have been used throughout the articles, to make it appear more human to the detectors, while it was largely generated by AI. Thus, AI as a writing-aid tool is very effective in generating text of good quality, and offers researchers a means of saving a significant amount of time, allowing them to spend more time producing and interpreting high quality data. This prompts further consideration about the future role of scientific writing for researchers.

Furthermore, the risk of scientific fraud induced by these language models is the cause of significant concern [15]. Generative AI has already been identified as an important cause of fraud, particularly in relation to the manufacturing of fake images [16,17]. Language models can facilitate fraud, particularly by creating false data. However, even in the presence of very efficient AI detectors, these tools evaluate the text, and not the data. Thus, if AI facilitates this type of fraud by becoming undetectable, it becomes imperative to look more specifically at the origin and quality of the data, as well as at the reproducibility of the results, particularly at a time where fraud and the retraction of articles is suspected to increase [18]. This also involves using a critical approach to examine scientific careers, which are often judged by the volume of articles published (publish or perish) rather than on the methodological rigor, the quality of the data, or the originality of the work.

Our work has several limitations. The main limitation is that we were not able to test all the detectors, and in particular Turnitin, which is considered to be one of the most effective currently available, because authorization is required in order to register. Other techniques for evading AI detection have been described but were

not used in this study, and it is likely that other techniques exist, as found using a simple Google search. However, we were still able to test 11 detectors but this rapidly evolving field does not allow exhaustive testing, as exemplified by OpenAI's recent abandonment of the development of their AI detector, which occurred during the review process of this article [19]. As the field of AI is evolving rapidly, the short or medium term relevance of our results may therefore be affected. In fact, other types of language models exist such as LaMDA developed by Google, which can be used by the Bard conversational agent (<https://bard.google.com/>). We chose to use a non-medical text due to the limitations of ChatGPT since the length of the generated text was prioritized to be appropriately investigated by these tools (i.e. short text are not adequately assessed, as well as extremely long ones since some tools have limitations regarding the text length). Similarly, we selected assessment of techniques for evading AI detection over the specific content of the text.

5. Conclusion

We observed that a text generated by GPT-4 can already evade most detectors, in a non-modified state. The strategies for minimizing AI detection are very effective and allow the text generated by the AI to pass for a human text, with all the detectors tested. In the absence of more efficient AI detection tools, and with the future improvement of these language models, there is a significant risk of no longer being able to determine whether a text is of human or AI origin, even without the use of artificial intelligence to minimize AI detection.

Disclosure of interest

D.J.Y. Yoon declares financial support from training with Guerbet and Coloplast Corin. G. Odri declares that he has no competing interest. Outside of this study, he has obtained research funding from the Fondation de l'avenir and received financial support from Sanofi Aventis and Stryker France, Johnson & Johnson, Medacta and Corin.

Funding

None.

Authors contribution

Guillaume-Anthony Odri and Diane Ji Yun Yoon designed the study, wrote and corrected the work in equal contributions.

AI usage statement

During the preparation of this work, the authors used GPT4.0 through the ChatGPT interface (OpenAI) to generate a text to be studied, to write the abstract, introduction and a part of the discussion sections. QuillBot (Course Hero, <https://quillbot.com/>) was used to paraphrase the text studied. Cognibypass (<https://cognibypass.com/>) was used to transform the text studied. After using these tools, the authors reviewed and edited the content as required and the authors take full responsibility for the content of the publication.

References

- [1] Siegel MG, Brand JC, Rossi MJ, Lubowitz JH. "Publish or perish" promotes medical literature quantity over quality. *Arthroscopy* 2018;34:2941–2.
- [2] Else H. Abstracts written by ChatGPT fool scientists. *Nature* 2023;613:423.

- [3] Elali FR, Rachid LN. AI-generated research paper fabrication and plagiarism in the scientific community. *Patterns* (N Y) 2023;4:100706.
- [4] Anderson N, Belavy DL, Perle SM, Hendricks S, Hespanhol L, Verhagen E, et al. AI did not write this manuscript or did it? Can we trick the AI text detector into generated texts? The potential future of ChatGPT and AI in sports & exercise medicine manuscript generation. *BMJ Open Sport Exerc Med* 2023;9:e001568.
- [5] Gao CA, Howard FM, Markov NS, Dyer EC, Ramesh S, Luo Y, et al. Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. *NPJ Digit Med* 2023;6:75.
- [6] Caulfield J. (2023, June 21). How do AI detectors work? Methods & reliability. Scribbr. Retrieved August 30, 2023, from <https://www.scribbr.com/ai-tools/how-do-ai-detectors-work/>.
- [7] <https://successfulstudent.org/how-to-make-chat-gpt-content-pass-ai-detection/>. [Consulté en juillet 2023].
- [8] <https://thenaturehero.com/bypass-turnitin-ai-detection/>. Consulté en juillet 2023.
- [9] <https://goldpenguin.org/blog/avoiding-ai-detection-for-chatgpt-writing/#:~:text=The%20easiest%20way%20to%20try,likely%20a%20computer%20wrote%20it.> Consulté en juillet 2023.
- [10] <https://brendanaw.com/how-to-make-ai-content-undetectable>. Consulté en juillet 2023.
- [11] <https://openai.com/blog/chatgpt>. Consulté en juillet 2023.
- [12] Charnley J. Surgery of the hip-joint: present and future developments. *Br Med J* 1960;1:821–6 [consulté en juillet 2023].
- [13] Weber-Wulff D, Anohina-Naumeca A, Bjelobaba S, Foltýnek T, Guerrero-Dib J, Popoola O, et al. Testing of Detection Tools for AI-Generated Text. *ArXiv:2306.15666*. doi:10.48550/arXiv.2306.15666.
- [14] Chakraborty S, Bedi AS, Zhu S, An B, Manocha D, Huang F. On the Possibilities of AI-Generated Text Detection. *arXiv:2304.04736*. doi:10.48550/arXiv.2304.04736.
- [15] Majovsky M, Cerny M, Kasal M, Komarc M, Netuka D. Artificial intelligence can generate fraudulent but authentic-looking scientific medical articles: Pandora's box has been opened. *J Med Internet Res* 2023;25:e46924.
- [16] Bik EM, Casadevall A, Fang FC. The prevalence of inappropriate image duplication in biomedical research publications. *mBio* 2016;7:e00809–16.
- [17] Gu J, Wang X, Li C, Zhao J, Fu W, Liang G, et al. AI-enabled image fraud in scientific publications. *Patterns* (N Y) 2022;3:100511.
- [18] Gaudino M, Robinson NB, Audisio K, Rahouma M, Benedetto U, Kurlansky P, et al. Trends and characteristics of retracted articles in the biomedical literature, 1971 to 2020. *JAMA Intern Med* 2021;181:1118–21.
- [19] <https://openai.com/blog/new-ai-classifier-for-indicating-ai-written-text>.