Summarizing scientific publications: integrating AI into our daily work

Niina Nuottamo[a], Rosie Morland[b], Dan Smith[b] and Remon van den Broek[a]
[a]Excerpta Medica, Amsterdam, Netherlands; [b]Excerpta Medica, London, UK

Scan the QR code or **click here** for additional interactive poster content

OBJECTIVES

The objective of this analysis was to compare currently available artificial intelligence (AI) tools versus medical writers (MWs) in their ability to create summaries of scientific publications, specifically looking at time-efficiency, readability and quality.

RESULTS

TIME

Panel A: Time required to write 1 summary (200 words)

MW 118 minutes

AI 2.4 minutes

Panel B: Time required to write 5 summaries

MW 588 minutes

AI 12 minutes

Figure 1: Mean time required for medical writers and AI tools to produce one (panel A) and five (panel B) summaries. Each summary was 200 words in length.

EASY TO READ

MW 3.6
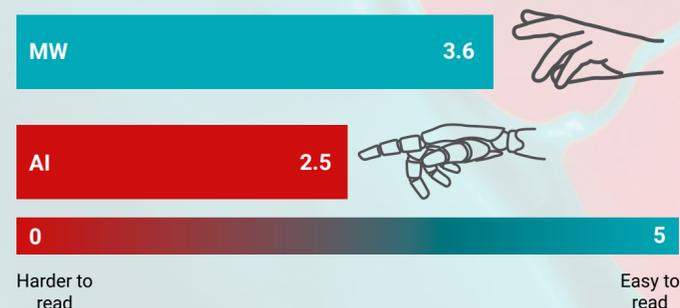
AI 2.5

0 — 5
Harder to read — Easy to read

Figure 2: Subjective assessment of readability of summaries generated by MWs and AI tools. Reviewers were posed the question "How easy the summary was to read" and asked to rate each summary using a scale of 0−5, where 5 was the most "easy to read". Means are reported, based on N=10 reviewers

QUALITY

MW 1.7

AI 3.4

0 — 5
I would send this to client — I would NOT send this to client

Figure 3: Subjective assessment of quality of summaries generated by MWs and AI tools. Reviewers were posed the question "How much work does this require before you would feel comfortable sending it to a client?" and asked to rate each summary using a scale of 0−5 where 5 was the most "work required". Means are reported, based on N=10 reviewers

DISTINGUISHING FEATURES

- Flow 21%
- Data included 21%
- Formatting 18%
- Sentence contruction 16%
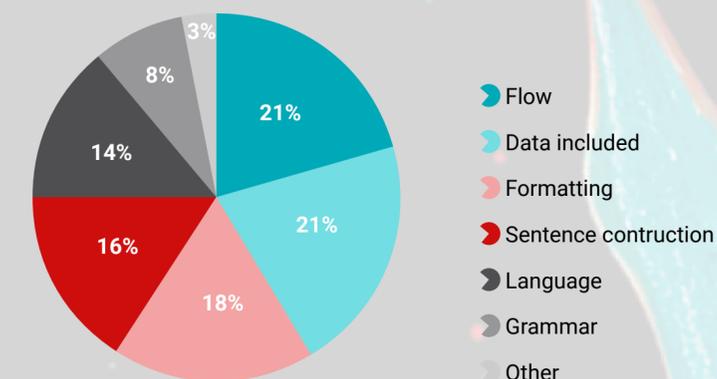- Language 14%
- Grammar 8%
- Other 3%

Figure 4: Reviewers were posed the question "What factors indicated to you whether the summary was written by a human or an AI tool" and given drop-down menu options. These are percentages for each answer based on N=10 reviewers

METHODS

- The body text (excluding abstract) of five scientific publications were inputted into two, publicly available, NLP-based summary tools.

- Concurrently, two MWs were also tasked to write unstructured summaries.

- Ten reviewers were block-assigned five batches of blinded summaries, comprising of four samples (2 AI-generated, 2 MW-generated) and scored their readability and quality using 5-point Likert-scale questions.

- Descriptive statistics were used.

CONCLUSIONS

AI tools were more time-efficient, while summaries written by MWs were easier to read, had better flow and included more relevant information.